

Telling the truth on yellow rain

The State Department is sticking to its allegation that yellow rain in South-East Asia is proof of Soviet use by proxy of biological weapons. To pay attention to other explanations would be more prudent.

PROFESSOR Matthew Meselson of Harvard University made a brave attempt last week to place the yellow rain controversy on a scientific footing. Judging by the swift — and disparaging — reaction from the US State Department, he appears to have failed. The loser, sad to say, is the State Department itself. What Meselson did was to make public the suggestion that yellow rain may have a natural explanation. In so doing, he challenged the State Department to substitute reason for faith in its contention that yellow rain is a toxin weapon manufactured by the Soviet Union and used by Soviet allies against resistance fighters in South-East Asia and Afghanistan.

The State Department apparently chose faith. For more than a year now, it has maintained that intelligence reports, refugees' stories and — most persuasively — chemical analyses point to Soviet use of toxin weapons. Analyses of environmental samples, whose specific origins have never been adequately documented, show the presence of fungal toxins known as tricothecene mycotoxins — "unequivocal evidence", the State Department says, of weapons use. If the State Department is right, it means that the Soviet Union is guilty of violation not only of humanitarian standards but also of two treaties forbidding the use of chemical warfare or even the possession of toxin weapons.

Meselson's theory begs the State Department to take a more critical look at its "unequivocal evidence"; the State Department's failure to do so only casts new doubts on its credibility. Meselson's idea, which he presented publicly at last week's meeting of the American Association for the Advancement of Science (and which was reported in *Nature* 5 May, p.3) is that the yellow spots containing tricothecene found on leaves and rocks from South-East Asia are at least initially the work of bees. Bees excrete large amounts of pollen when they leave the hive and Meselson has even collected such samples from a parking lot at Harvard. He has been led to his hypothesis that yellow rain may be pollen-laden bees' excrement by several reports in the past year of the discovery of pollen in samples of yellow rain. Nobody says that this hypothesis rules out the toxin weapon hypothesis, and indeed the occurrence of mycotoxins in samples of yellow rain remains to be explained. One possibility is that natural bees' excrement has been contaminated by weapons toxin, another that the material is a natural substrate for the growth of naturally occurring *Fusarium* fungi. Meselson takes no position on these alternatives. His demand is merely that all explanations of the phenomenon of yellow rain should be seriously investigated.

Argument

The State Department's response so far has been shabby — to dismiss Meselson's argument as the "great bee caper". It must be particularly embarrassing for the State Department that Meselson, a respected student of all governments' policies on chemical and biological weapons, should be the one now seeking to help it understand how science should be done. The department has responded in the same way to other evidence recently brought to light: either the evidence has been dismissed or it has been welcomed as further proof of Soviet treachery. The first reports of pollen in yellow rain were thus greeted with the prejudiced opinion that pollen must have been deliberately added to make toxins more easily inhaled and so more effective. Last week's comment on Meselson's argument that there may be a

natural explanation of yellow rain shows that prejudice can lead to blindness: the State Department's opinion that a natural explanation is ruled out because the quantities of toxin found in yellow rain would be lethal to bees entirely misses the point that, on Meselson's hypothesis, toxins would be found (or accumulated) only after the bees had excreted the pollen.

Evidence

In presenting his theory last week, Meselson went through the exercise of listing where the evidence supports his theory and where it contradicts it. The State Department seems not to have taken the hint; nor did it admit that it should define the areas where questions remain. But there is a host of unanswered questions. How widespread is the phenomenon? Does it occur in parts of South-East Asia remote from battle areas? Which *Fusarium* species are native to the region? Does the pollen in the samples vary from region to region? From season to season? Is it consistent with the vegetation of the region? Why is there such a diversity of pollen in the samples found? Do bees feed on such a diversity of plants? These are the sorts of questions the State Department should be asking if it is to preserve any credibility on this issue.

The tragedy is that the State Department has impaled itself on a hook from which it cannot escape by making public serious charges of Soviet perfidy before these could have been sustained by evidence. Politically, it would not be acutely embarrassing to have to admit that some of the data are in dispute. By discarding objectivity, the State Department has now laid itself open to the charge that it is more anxious to sustain its publicity campaign on yellow rain than to search for the truth. Although Meselson's hypothesis was communicated to government officials more than a month ago (*Nature*, *ibid.*), the department threw together its in-temperate response only when he made the notion public. To judge from what State Department spokesmen have been saying, Meselson's hypothesis is regarded as a threat, not as a legitimate question.

The State Department is also putting out a line that has become increasingly apparent as new questions have arisen about the scientific data — the assertion that the case that yellow rain is a toxin weapon rests not only on scientific evidence but on intelligence data. This is disingenuous. The State Department knows as well as anybody that such credence as its allegations against the Soviet Union have so far won derives from the way in which anecdotal evidence from South-East Asia has been fortified by what purports to be objective evidence. The United Nations investigation in South-East Asia and that mounted separately by the Canadian Government have underlined the difficulty of making sense of what is said (often in an obscure dialect of a foreign language) by politically displaced persons sheltering in refugee camps. *The Yellow Rainmakers*, a book by Grant Evans, an Australian sociologist (Verso Editions), provides an explanation of these difficulties — that of the peculiar social structure of the Hmong population against whom the present government of Laos has taken up embittered cudgels. The State Department cannot now rest its case on this flimsy evidence without, in effect, acknowledging that the allegations originally put forward in 1981 by then Secretary of State Alexander Haig, but repeated last year by his successor Mr George Shultz, appear increasingly insubstantial. [.]



Thatcher's government

The British Conservative Government is almost certain to be reelected but should change.

A FEW days from now, Mrs Margaret Thatcher, British Prime Minister for the past four years, will have returned to Downing Street and will probably be too busy to remember the curious general election campaign that has occupied the British for the past month. It is, however, less certain that her ministerial colleagues will slip back as easily into their old routines. Some will no doubt have lost their parliamentary seats in the election today (9 June). Others will have been posted elsewhere in the government, while still others will have lost their jobs altogether, condemned (for the time being) to the life of an ordinary parliamentarian with no black limousine in attendance. That is a fate that many in the previous government deserve.

The reasons for the return of the Conservative Government, earlier this week apparently a foregone conclusion, will be endlessly debated, but the fact of it is remarkable. With more than three million people now unemployed, with industrial production no greater than fifteen years ago and with all kinds of social institutions (universities for example) fearful for their future, the government's electoral chances should have been slim. Moreover, unlike the other parties competing in the past few weeks for votes, the government has made no promises of improvement. More of the same has been offered as a recipe for the next four or five years. In at least three respects where the scientific community is concerned, however, the recipe needs radically to be changed.

Management

First, the administration of public support for the research enterprise needs to be overhauled. During its previous spell in office, the British Government resisted the argument (from the House of Lords Select Committee on Science and Technology) that there should be some means by which government expenditure on research and development could be more effectively assessed and coordinated, perhaps by saddling a minister with part-time responsibility for the task. The proposal was turned down, ostensibly on the grounds that the Prime Minister, herself trained as a chemist, was well placed to carry out that task, but more probably because powerful government departments believed the present chaos (also called plurality) more congenial. The results, however, have not been creditable. Little has been done to repair the breakdown of the dual-support system by means of which university research is supported, defence research remains virtually a law unto itself, while government policy has been driven by the fashionable pre-occupations of some of its ministers — the widespread conviction that the universities are effete and, contradictorily, excesses such as the love affair with what is called information technology.

The flaw in the present arrangements, more obvious when the pattern of public support for research and development has been changing fast, is that there is no mechanism by which innovations can be intelligently discussed, and their consequences guessed at. The result is that when an influential cry goes up for more of this or less of that, every government agency does the same thing. The remedy is simple — to separate the research councils from education, to replace the Advisory Council on Applied Research and Development (a topic-oriented body) by a mechanism for reviewing the work of all the public agencies in research and development, and to ensure that they know individually what they are for. A part-time minister should indeed be put in charge.

Second, government administration of higher education should be made more responsible. The universities appear to have survived the budget reductions of the past two years, not always deservedly. But they are still unclear what the future holds for them. When and how the University Grants Committee will lift the invidious quota system which at present prevents universities from competing among themselves for students, and thus in efficiency, remains to be decided. Meanwhile, perhaps only temporarily, the gulf between the university system and the poly-

technics has if anything been widened.

For several years, the overriding difficulty in these arrangements has been the fitfulness of government policy. Budgets have been cut and then partially restored. Even imaginative administrative innovations (such as the scheme for running a handful of universities on cash budgets and a loose rein — see *Nature* 26 May, p.270) have been canvassed in a way that unsettles the generality of the government's dependants. Nobody should object to change, but hasty unplanned change serves no purpose. Mrs Thatcher's new government needs somebody in charge whose policies will be more steadfast.

School education is a still more formidable problem, largely because the government does not appreciate what it consists of. Irrelevantly, the previous government was much concerned with schemes for broadening parents' choice of which secondary schools their children should attend. Rightly, it was concerned with the improvement of "standards" in secondary education. Wrongly, it has not grappled with the central flaw in British secondary education that students, especially the most able, are educationally crippled by the narrow curriculum they must follow on the way to higher education. The government's election manifesto boasts of steps already taken to augment technical education in secondary schools. How many of those who profit will be able to speak some language other than English? Nobody objects to what the manifesto calls the pursuit of excellence in the schools, but so far it has been conducted too narrowly.

Third is the long-standing British problem, the economy. The government has insisted throughout the general election campaign that it cannot properly have a direct role in industrial innovation but that its task is merely to create the social and economic circumstances in which new enterprises can flourish. The principle is impeccable, but has not been followed in the past four years and may not be in the next five. In spite of having committed £2,000 million during this year to the training of young people who would otherwise be out of work, the British Government is spending less on vocational training than its industrial competitors (according to a document by the National Economic Development Office leaked to the Labour Party early in the election campaign). Fees paid by individuals to educational institutions must still come from taxed income. Spending on research and development is substantial when measured as a percentage of national income, but the gross domestic product is growing only slowly, while the absence of a mechanism for making research policy coherent means that much of what is spent is irrelevant to the government's declared purpose of encouraging innovation. Meanwhile, for reasons which are politically understandable but economically nonsensical, the government has continued to subsidize loss-making nationalized industries while preventing (in the interests of not borrowing the money) profitable state industries from becoming more efficient by the investment of sufficient capital. And the government's declared belief in competition as a spur to efficiency is belied by its attempt (cut short by the election) to sell off the country's chief telecommunications network (now nationalized) with entirely inadequate restrictions of its monopoly power.

Economy

Nobody pretends that the British economic climate can be changed by the Department of Industry acting on its own, but it is the only government department with a remit to intervene directly in industrial affairs. While its behaviour in the past four years has been in many ways consistent with what the government has been saying, the department has been too much constrained by timidity from using its spending power to make nationalized industries efficient and has also been the chief source of untrustworthy fashion, that for information technology in particular. A few months from now, the promise in the government's manifesto that its plan to sanction new cable television networks will be a substantial contribution to the emergence of new technology will seem ridiculous. The trouble seems to be that the government cannot easily tell how its principles should be applied in practice. It needs a clarity of purpose that it has so far lacked.

||

French universities

Government makes only token concessions to Right

WITH most French students now sitting examinations in preference to demonstrating on the streets, debate on the government's new bill to reform higher education was a little more ordered last week — but only a little. The National Assembly is discussing the bill clause by clause, but at the pace of an *escargot*, because of the enormous number of amendments tabled by the opposition: some 1,950. This is more than the number faced even by the hotly argued nationalization bill or any other bill prepared under M. Mitterrand's presidency. There are sub-amendments to sub-amendments which add little substance to the argument of the amendment. The first clause (of 68 in the bill) took nearly eight hours to be approved, and there are projections that parliament will need a special session in July to see the bill through just the lower house of parliament, while other important legislation remains undebated. The opposition parties are being accused of filibustering, sabotage and devaluing democracy. Clearly, they sense the kill. De Gaulle fell after the university rebellion of spring 1968; will Mitterrand go the same way? The Right hopes so.

However, M. Alain Savary, the minister for education, is doggedly resisting all attacks, and yielding only a small amount of ground. Ultimately, most of the amendments proposed by the opposition are being rejected, and something very similar to the minister's first intentions is emerging from the mill.

Except, that is, on one point. Savary wanted greater control of the *grandes écoles*, the engineering schools which seem unfailingly to train the elite of the civil service, government and industry. Most of these schools belong to other ministries: the Ecole Polytechnique, for example, arguably top of the heap, is controlled by the ministry of defence. Article 9 of the bill indicated that rules developed for institutions under Savary's direct control — such as universities — could be extended "by decree" to other institutions. The *grandes écoles* and the Right objected, and Savary bowed: any extension of the rules would be made only with the "good advice" of the schools themselves and of their ministries, he said. This gave a "minimum guarantee", the opposition felt, and Article 9 was thus amended. So the bill, although nominally about the whole of higher education, becomes predominantly one about universities.

On most other matters, however, the government has won the day. Universities will be run not just by two councils, as at present (administrative and scientific) but, as the bill outlines, by three. The third, for

"studies and university life", will focus on the students and "guarantee their political and union freedom". Above the councils will come the president of the university (effectively the vice-chancellor), who must be a practising researcher/teacher and who will be elected. The electorate will include not just the professors but the whole teaching staff (provided they are permanent and French). This, said one socialist in the Assembly, would put an end to the domination of the universities by the "mandarins" (the professors); but according to a member of the opposition, it would open up the presidency to political and union lobbying, thus further politicizing the already political universities.

As for examinations for students to pass from the first stage of university (the first one or two years) to the second — a proposal that stimulated much of the student protest — the text of the bill has merely been made more precise. In general, there will not be such examinations; but they will be imposed when the second stage would otherwise be overcrowded or when

ultimate employment for all students who would take the course seems unlikely.

And as for increased autonomy — pressed for by the opposition despite opposing it when in power — the universities will see little of it. One amendment would have given the universities the right to award their own degrees, rather than the national diplomas presently awarded. A socialist deputy responded that the opposition clearly wanted "the North American system", and would end up privatizing the universities and requiring students to pay their own fees. The amendment was rejected.

The Savary bill will open the first stage of universities to more students (the Right say another 300,000; Savary says 10,000 per year, or 40,000 in all); it may (or may not) turn more away at the entry to the second stage. It will turn the third stage into something close to a PhD — an idea welcomed by most French scientists. It will turn the universities' attention more towards industry and the jobs their students will enter. It will decrease the power of the professors over "their own" institutions, and increase the decision making structure within the university. It will leave the universities still broadly centrally managed and it will leave the *grandes écoles* largely alone. And it will leave France as divided as ever about the proper role of its universities.

Robert Walgate

Synchrotron radiation

European rivalry heats up

Hamburg

AFTER the frustration of a prolonged shutdown in 1982 and a period of unreliability, the synchrotron radiation is now so good at DESY in Hamburg that the European Molecular Biology Laboratory (EMBL)'s outstation there can hardly keep up with the flow of visitors using its facilities. More staff have been budgeted for but have yet to be recruited. Without them — perhaps, in any case — the facilities at EMBL may become second choice for those with access to the new facilities at Daresbury in the United Kingdom operated by the British Science and Engineering Research Council.

In theory the staff of the EMBL outstation should be growing in number thanks to an extra budgetary allocation won for them by Dr Lennart Philipson, director general of EMBL in Heidelberg. The problem is that since then Dr Juan Bordas, who co-directed the outstation with Dr Michael Koch and has just moved to Daresbury, and two other staff members have moved on without being replaced. For now, the remaining staff have little time to carry out their own research and are up to their necks in assisting visitors.

Visitors come from many European countries, Israel and the United States. On the current list of 60 projects accepted by the priorities committee, one-third come from the Federal Republic of Germany and

one-sixth from the United Kingdom. The next most represented countries are the United States (6 projects) and Sweden (5). About half of the projects are straightforward protein crystallography. Most of the rest involve time-resolved scattering measurements, particularly on muscle and muscle related proteins.

As to the immediate future, the EMBL outstation knows it will increasingly feel competition from Daresbury now that the latter is really in operation, two years behind schedule. The one great advantage of Daresbury over DESY is that it is totally dedicated to synchrotron radiation. At DESY, priority is given at certain times to the colliding beam experiments of high-energy physicists. At such times the parasitic use of synchrotron radiation can continue but it is less than ideal.

That apart, opinions are mixed about the relative merits of the two facilities and depend to some extent on individual experiences at DESY and expectations of Daresbury. One view is that the lack of organization is bound to count against DESY. Another is that Daresbury has a long way to go before it can offer such good facilities as DESY for those who wish to carry out low-temperature work or time-resolved studies, for both of which the instrumentation developed by the EMBL staff has been invaluable. Peter Newmark

British nuclear weapons tests

Australians wait and see

Canberra

A GOVERNMENT report on the effects of the British nuclear weapons tests in Australia during the 1950s has found them to be safe, in line with official statements from the British Ministry of Defence. The long-awaited Australian Ionizing Radiation Advisory Committee (AIRAC) review of operational safety measures and possible after-effects was tabled in the House of Representatives on 26 May. That the government is not altogether satisfied is suggested by its decision to await the results of a health survey of all servicemen involved in the tests (see *Nature* 300, 98; 1982) before deciding what to do.

The main conclusions of the AIRAC report, based on official records, are:

- There is no evidence of any departures from compliance with International Commission on Radiological Protection (ICRP) recommendations.
- No Australian received a radiation dose in excess of the then ICRP recommendations.
- Criteria for safe firing were met.
- There is no evidence that aborigines were injured by tests.

Even Totem 1, the first of the two tests at Emu, which was fired before either the UK Atomic Weapons Research Establishment at Aldermaston had issued safety regulations or the Australian Safety Committee had been formed, and whose plume is incriminated as the "black mist" of aboriginal lore, was safe, the committee concludes. It was after tracking the clouds of Totem 1, too, that aircraft contaminated with radioactivity were not isolated or treated for at least five days. Totem 1, the report concedes, may have exceeded current ICRP recommendations but "if that limitation was in fact exceeded, the excess would have been small and there would be no detectable effect on persons". Moreover, "a limited number of air crew may have been exposed to transient concentrations of radioactive substances exceeding the derived levels recommended for continuous exposure over a 13-week period, but not to total radiation exposures in excess of the recommended limits".

The conclusions relating to radiation exposure were based on records of readings from personal radiation monitoring devices. It has, however, been alleged in the past few months that dosimeters were faulty, that some personnel were not issued with film badges, that badges were issued but not processed, that some personnel were lax about wearing them and that some badges turned black and could not be read. In these circumstances, the recorded levels would not be relevant, especially because in the 1950s the policy was not to issue badges unless "specific exposure was likely to be a substantial fraction of prescribed limits".

The aboriginal claim that a "black mist"

at the time of Totem 1 led to diarrhoea and death was dismissed by AIRAC on the grounds that the stories were inconsistent, perhaps provoked by patrol officers explaining the nature of nuclear tests. The committee also noted that there was a contemporary measles epidemic.

But the issue of nuclear tests in the Southern Hemisphere is still controversial.

Australia has expressed strong indignation at the second of this year's round of French nuclear tests in the tiny atoll of Mururoa in the South Pacific. What the government had been led to believe would be a "trigger device" turned out to be a 70 kilotonne explosion. The last Labor government successfully prosecuted France at the International Court of Justice, but the tests merely went underground. Mr Bob Hawke, the Prime Minister, is expected to protest in person when he meets President François Mitterrand of France in Paris later this month.

Vimala Sarma

American Geophysical Union

Are nuclear test bans verifiable?

Baltimore

A SERIES of talks and a panel discussion at last week's meeting of the American Geophysical Union (AGU) provided a cross-section of seismological opinion on two central issues relating to a possible Comprehensive Test Ban Treaty (CTBT). First, have the United States and Soviet Union complied with the 150 kT threshold imposed by the Threshold Test Ban Treaty in 1976? The answer seemed to be: yes for the United States, impossible to say for the Soviet Union given uncertainties in the seismological evidence. Second, would a CTBT be seismologically verifiable? The answer depends not only on the ability to discriminate between earthquakes and explosions but on the presence of monitoring arrays in the Soviet Union.

The AGU discussion was prompted by the Reagan Administration's indifference to current test ban negotiations. All US and Soviet tests have been conducted underground since the limited Test Ban Treaty was ratified in 1963. Although the threshold treaty was not ratified, the United States and Soviet Union undertook to abide by it. The United States, the Soviet Union and the United Kingdom also discussed a comprehensive test ban, but the Carter Administration withdrew in 1980 following the Soviet incursion into Afghanistan. The United States now argues that (1) seismic verification techniques are insufficiently reliable; (2) the Soviet Union may have repeatedly exceeded the 150 kT threshold since 1976; (3) there is a need for continued testing.

The most pessimistic views were expressed by Drs T. C. Bache and R. W. Alewine of the Defense Advanced Research Projects Agency. Dr Alewine, assessing seismic estimates of Soviet test yields, concluded that the evidence was consistent with Soviet violation of the 150 kT threshold on several occasions since 1976. He acknowledged that a geophysical explanation is possible, but emphasized that inoperability of treaty protocols on data exchange did not permit investigation of that possibility. But his estimates were challenged by other speakers. For instance, Professor Lynn Sykes (Lamont-Doherty Geological Observatory) said that too

much attention had been given to body waves (*P* waves) and that surface (Rayleigh or *S*) waves, the attenuation of which is less variable from one region to another, indicate Soviet compliance with the threshold treaty.

Sykes and others described empirical relationships between surface and body wave magnitudes on the one hand and test yields on the other using data from tests at the Nevada test site. A crucial element in the dispute is the correction for body wave attenuation needed when using such relations for Soviet tests at the site at Semipalatinsk. During the panel discussion, E. Herrin (South Methodist University) claimed that, given the uncertainties, one could say with conviction only that maximum Soviet yields were greater than 75 kT and less than 300 kT.

Body-wave corrections are also important in discriminating between explosions and earthquakes — essential of test ban verification. Underground explosions produce a larger proportion of seismic energy in the form of body waves (as opposed to surface waves) than do earthquakes. However, the attenuation correction for body waves needs to be relatively large (a few tenths of a magnitude) to ensure that earthquakes and explosions form distinct statistical populations within this criterion. T. W. Bache emphasized such problems and maintained that they, together with practical problems of assessing huge volumes of data, render verification impossible. Other participants (particularly J. Evernden of the US Geological Survey) attacked Bache's adherence to a zero correction for body-wave attenuation.

Sykes mentioned a difficulty associated with the lowering of the test threshold. 150 kT devices are exploded at depths that ensure a hard rock surrounding. But small explosions could be fired above the water table in soft rock, which would reduce by an indeterminate amount the energy coupled into seismic waves. However, Evernden surveyed the uses of depth, location and frequency spectra as seismic discriminators and expressed the view that, provided that non-seismic monitoring techniques were also available, discrimination was always possible. **Philip Campbell**

Information technology

Europe's bid to catch up

Brussels

THE European Commission's research officials holding their breath as the time approaches for launching the Community's most ambitious research project yet — the Esprit programme of research and development in information technology. By the deadline for the submission of proposals late in February, nearly 200 specific proposals had come from all member states with a total combined cost of 110 million European Currency Units (ECU) (1 ECU = £0.55). Member governments will consider the proposals at a research council meeting later this month and again at a meeting in October.

The immediate issue is the "pilot phase" of the Esprit programme, when there will be 15 one-year research projects ranging from advanced microelectronics to software technology, advanced information processing, office automation and computer integrated manufacturing. A sixteenth project will be aimed at improving the exchange of information between the participants.

Commission officials note that applications for the pilot phase come from all member states and that 620 separate organizations are involved (some of them more than once) in the applications. The Commission is hoping that the outcome of the pilot phase will persuade governments to give the go-ahead for the first of two five-year operational phases at a cost of 1,500 million ECU.

The Commission is confident that it will succeed, partly because the project was first suggested in 1982 by the 12 principal electronics companies in the Community and partly because the pilot phase and the subsequent steps have been planned in close collaboration with about a hundred speculators from industry, universities and research institutes in Europe. Earlier this year, the electronics companies told the Commissioner for Research and Industry, Vicomte Etienne Davignon, that unless a cooperative industrial programme could be mounted on a sufficient scale, "most if not all of the information technology industry could disappear in a few years".

Financially, the electronics industry is prepared to back the project with its own funds. It will provide at least 50 per cent of the total cost of Esprit research and development, not less than 1,500 million ECU for the 10-year operational phase and 11.5 million for the pilot phase.

Even the pinchpenny governments of the Community agree on the need for action. The Commission calculates that while the Community's trade balance in information technology was level in 1975, it showed a \$5,000 million deficit in 1981 and twice as much in 1982. Now, the Commission says, Community-based industry supplies less

than half of the European market in information technology, which itself accounts for 34 per cent of a world market estimated to be worth \$237,000 million in 1980. With the world market expected to reach \$500,000 million a year by 1990 (at 1980 prices), the Commission thinks that "the identity of Europe and its political independence" is at stake.

At the same time, the Commission admits the Esprit programme is small compared with the present rate of research and development expenditure in Europe in information technology, estimated at \$5,000 million a year, and with the investment of the largest US companies, now spending individually \$2,000 million a year on research and development. The Commission also notes that corporate joint ventures such as the semiconductor research cooperative have been actively encouraged by the Reagan Administration's Economic Recovery Tax Act. And, inevitably, the Japanese investment of \$500 million on a fifth-generation computer programme is a spur.

Even so, Davignon considers that Esprit will stimulate strategic thinking and a growth of confidence. The Commission's view is that in Europe there is insufficient multidisciplinary research, too large a gap between university and industrial research and a mismatch between the skills of university graduates and the needs of industry.

European industry, while welcoming the Esprit programme, has a number of particular concerns. Thus a source at Philips (Netherlands) emphasizes the need for access to research results and good management on the programme. The Commission has decided that the ownership of the right to exploit new information will normally rest with its contractors but that teams working on other Esprit projects should have privileged access to research results.

On management, the Commission plans a management committee (representing member states), consultations with industry and also to use the pilot phase as a means of testing different ways of managing the full programme.

The big question still undecided is whether governments will in the end agree to contribute 1,500 million ECU — a sum expected to come as something of a surprise some of the larger member states with extensive research programmes of their own. There may also be difficulties about the Commission's plan to assume overall responsibility for the programme, depriving governments of their supervisory powers. But one Commission official said that "time is running out" and that the scale of the programme "merely reflects the extent to which we lag behind".

Geert Linnebank

UK computing

Half-hearted sponsorship?

Mr Brian Oakley, who took charge of the Department of Industry's new unit for promoting computer developments on 1 June, stresses that the directorate will adopt a flexible approach in deciding what areas of advanced information technology to support. The new unit, known as the Alvey directorate, will coordinate a £350 million research programme. Some £50 million of this will be spent in academic institutions, but some features of the scheme have been criticized by the industry.

Mr Patrick Jenkin, Secretary of State for Industry, announced the government's response to the Alvey report at the end of April (see *Nature* 5 May, p.2). The Alvey committee, set up by the Department of Industry to ponder how to help British industry to compete successfully in computer technology, had reported last year. The government agreed to set up the directorate with a 5-year programme to develop key areas in the field of advanced information technology, and provided a total of £200 million to support it. Industry will, it is hoped, provide a further £150 million.

The chief departure from Alvey's recommendations is that the government's contribution to any one research project will be limited to 50 per cent of the total cost. Alvey suggested 50-90 per cent.

Oakley says that requiring industry to bear half the cost of a project (as in the Esprit programme) ensures it "means business"; small companies should be able to keep up if they pool their efforts at the pre-competitive stage of development. He acknowledges the difficulty of encouraging cooperation, especially since there is a shortage of top experts in the industry. For this reason he is committed to improving education in information technology. The Alvey committee, of which Oakley was a member, pointed out that simply putting a microcomputer in every school will only breed a generation of poor BASIC programmers. Part of the government's contribution to the Alvey programme is earmarked for education but the details have yet to be settled. Oakley, previously secretary of the Science and Engineering Research Council, is determined that it will be used effectively.

Mr David Fairbairn, director of the National Computing Centre Limited, has welcomed the programme but expressed some reservations. He says that the 50 per cent maximum subsidy is adequate in fields where the first developer can expect the lion's share of the profits, but in areas which cannot easily be protected by patent (such as software development) the incentive may not be sufficient. Fairbairn also feels the programme fails to put enough weight on bringing technologies to the market.

Tim Beardsley

UK research careers

ARMS call for job security

PARTICIPANTS at a symposium held by the Association of Researchers in Medicine and Science (ARMS) last week heard vigorous calls for a planned career structure for science researchers in Britain. In a keynote speech, Dr J.P. Dickinson, the association's chairman, said that widespread lack of concern over the present system of research support, under which a high proportion of researchers are retained on short-term contracts, means that research workers are inadequately trained. ARMS was formed in July 1978 after an advertisement in *Nature* attracted world-wide support. Originally concerned with biomedical research, the association has now broadened its scope to cover all researchers in the United Kingdom. One of its objectives is that a higher proportion of researchers should be retained on a professional basis beyond the age of 35. This, the association says, could be achieved within the existing research budget by reducing the overall recruitment to research careers by 20 per cent.

ARMS admits that its proposals would reduce the number of research projects in progress but believes that, by employing senior and experienced staff with professional status, productivity would be increased. The case is backed up by a detailed model of a proposed career structure in medical research.

ARMS is at pains to make clear that it is not calling for life-long security of tenure for researchers, but rather what it calls "improved continuity". There should be a professional body to guard the interests of career researchers, and changes in the mechanism of research support are proposed that would retain the positive features of the present system (sensitivity to budget constraints) while ensuring that proper account is taken of overhead costs at research institutions. A unified administration would, according to Dr Dickinson, serve the interests of science by promoting interdisciplinary mobility.

Other participants echoed the need for mobility. Dr J.B. Wyngaarden, director of the US National Institutes of Health (NIH), stressed that the lack of an adequate career structure for biomedical researchers was an international problem, and pointed out how scholarship schemes run by NIH within a roughly level real budget have stimulated medical practitioners to enter research.

Professor W.S. Peart, of St Mary's Hospital, argued that clinical departments should make a practice of employing some non-clinical researchers to promote the exchange of scientific and medical views. Researchers should, he said, have more say in policy decisions.

Others were more concerned with the lot of young researchers and PhD students. Professor W.L. Ford, of the University of

Manchester, called for the "scandalously low" level of support for research students to be doubled, with a concomitant reduction in numbers. Other speakers urged the universities to grant full academic status to all their researchers and

to bring an end to the practice of insisting on waiver clauses in contracts which deny employment protection rights to those who sign them.

ARMS proposals would require radical changes in policy by research councils and universities. But the stress it is placing on financial realism and the benefits to science should ensure that it is at least listened to.

Tim Beardsley

University of Oxford

Selection process ripe for reform

PROPOSALS for simplifying the admission of undergraduates to the University of Oxford have been put forward by a committee under the chairmanship of Sir Kenneth Dover, president of Corpus Christi College. The major change proposed is that those applying to the university on the basis of their school-leaving "A-level" examinations should not in future take the special Oxford entrance examination. The committee says its proposals would probably increase the proportion of admissions from maintained (state) schools.

Oxford admissions policy has been controversial for two reasons (see *Nature* 24 March, p.277). The first is the charge of elitism: as the Dover committee puts it, "An Oxford college is still often perceived... as a club for decent chaps and well brought up girls". Last year, 47 per cent of admissions were from independent ("public") schools, attended by about 5 per cent of pupils in their age group. This imbalance reflects the high proportion of applicants from such schools: 39 per cent last year. Many maintained schools do not encourage their pupils to apply to Oxford.

The second source of controversy is the complexity of the admissions procedure. Since this was last rationalized in 1962, many of the colleges — which are constitutionally autonomous — have introduced reforms of their own, with the result that an applicant is now faced with a range of entrance options that varies between colleges and between subjects.

For some combinations of studies, the entrance examination must be taken, while for others entrance may be on the basis of an interview and reports; colleges have markedly different policies. This complexity is considered by some to deter potential applicants, especially those at schools without an established tradition of sending students to Oxford. The practice of letting post-A level applicants compete in the entrance examination with pre-A level applicants has also been criticized because it encourages early specialization and because many independent schools have expertise in post-A level coaching that maintained schools cannot provide. The system also creates difficulties for those trying to assess candidates' potential. The Dover committee was charged with finding

a simplified entrance procedure likely to be acceptable to all the colleges.

The thirteen college representatives on the committee acknowledge the "extraordinary complexity" of their task. As expected, a quota system to regulate the intake from different types of schools was rejected, on the grounds that this would inevitably prejudice academic standards. The committee says that pre-A level candidates should be free to choose whether or not to take the entrance examination, and that colleges should not in future set quotas for admissions with and without the examination. But the proposals contain nothing to prevent colleges preferring one or other mode of entrance from exercising their preferences; colleges are simply asked not to "push" applicants in one or other direction.

The Dover committee says that the practice of awarding scholarships to promising applicants before admission, on the basis of their performance in the entrance examination, should be ended. The awards, which are of little financial value, would in future be made only after admission. The present system has again been criticized for encouraging excessive specialization. Until recently, the main function of the awards was to allow less popular colleges to secure good applicants by offering scholarships which candidates were then obliged to take up. Many Cambridge colleges have already decided to abandon the scholarship system.

Dover's recommendations, though falling short of the radical proposals some had hoped for, will be popular with schools, especially maintained schools. Dr John Rae, headmaster of Westminster School (a leading independent school), says that the new scheme would ensure that justice is seen to be done. But it would be unwise to assume that all 29 of the colleges admitting undergraduates will find the recommendations acceptable, despite the committee's plea for unity. One likely consequence if the proposals were adopted would be an increase in the proportion of pre-A level candidates electing to take the entrance examination. Until now the promise or the hope of places reserved for non-examined candidates may have helped some colleges to attract promising students.

Tim Beardsley

Environmental Protection Agency

Pesticide approval challenged

Washington

As the new administrator of the Environmental Protection Agency (EPA), William Ruckelshaus, was ordering equal access for all interested groups to the agency's policy-making, the legacy of Anne Gorsuch's controversial reign was being challenged in court. In a suit filed by an environmental group and AFL-CIO (a federation of labour unions) late last month, the agency was accused of holding secret meetings with pesticide manufacturers at which key decisions on the approval and retesting of their products were made.

The suit, which will be heard in the federal district court here, seeks to invalidate decisions that the plaintiffs say ignored adverse safety data and to require a court-supervised review of all agency actions on pesticides taken over the past 18 months.

According to Jackie Warren of the Natural Resources Defense Council (NRDC), the environmental group involved in the suit, the agency held so-called "decision conferences" with the manufacturers under a policy adopted by Gorsuch and the former assistant administrator for pesticides, John Todhunter. Warren said the action violated statutory requirements for open hearings and public comment.

NRDC also charges that through these conferences, manufacturers were allowed to write entire sections of registration standards for certain chemicals. These standards review health and safety data for a chemical, and provide the framework for subsequent registrations of pesticides containing that chemical. A separate registration is required for each specific formulation and for each intended use of a pesticide.

The suit specifically mentions permethrin, a pesticide approved for use on cotton and for which manufacturers are seeking registration for use on edible crops; and two herbicides, picloram and glean.

Another target of the suit is EPA's decisions on whether to order in-depth reviews of data on already-registered pesticides. NRDC says that no reviews were ordered in fiscal year 1982, and further that several chemicals on a preliminary list for review were removed after meetings with manufacturers. NRDC says the agency has backed down on earlier proposals to restrict the use of Lindane, a pesticide, and pentachlorophenol, a wood preservative. EPA's decision on the latter, NRDC says, was announced after ten closed meetings with manufacturers; a single public meeting to receive comment on the decision was announced on the stationery of the wood preservation industry.

EPA counters that all registrations have been published as proposed rules, with an

opportunity for public comment before being made final. Jim Roelofs of the agency's office of pesticides said that discussion with manufacturers is "a way of doing business that seems reasonable to us — we do it on a daily basis". He added that

Australian animal virology

Politicians fear foot and mouth

Canberra

THE federal caucus (consisting of all Labor Party members of parliament) last month (24 May) banned the import of live foot and mouth disease virus (FMDV) into Australia for five years. The decision, expected to be endorsed by the cabinet, thus agrees with a recommendation by the Australian Science and Technology Council (ASTEC) that the new Australian National Animal Health Laboratory (ANAH) should not work with the live virus, the task for which it had been built at a cost of A\$150 million (see *Nature* 19 May, p.190).

Pending the cabinet decision, Dr Ken Ferguson, director of the Institute of Animal and Food Sciences of the Commonwealth Scientific and Industrial Research Organization (CSIRO), the body which will run ANAH, said that CSIRO would "broadly accept the recommendation". In any event, an internal review of the Division of Animal Health shows that CSIRO would not be in a position to handle live virulent virus until mid-1986, and the ASTEC recommendation delays this schedule only by about eighteen months.

The controversy about the import of

such conferences have been routine since long before the Reagan Administration.

Roelofs said the purpose of the meetings was mainly to establish what data the manufacturers would need to show in support of a particular registration. And while acknowledging that 5 of the 40 registration standards so far issued had been written in collaboration with manufacturers, he said that this was a limited "experiment". **Stephen Budiansky**

FMDV has riven the scientific community, at least in Canberra, and both sides derive tempered support from the caucus decision. Ferguson claims that the opponents of FMDV importation in fact begrudge the money being spent on the new laboratory, and that they have taken up the importation issue as a tactical manoeuvre to win community support after finding the expense argument politically unsaleable.

Another disputed issue is whether live virus would be needed for diagnosis should a putative outbreak occur. Ferguson says that people actually working with FMDV would not feel happy about making a diagnosis without live virus controls, particularly in the sensitive virus neutralization test. On the other hand, Dr Graeme Laver of the department of microbiology at John Curtin School of Medical Research, Australian National University, a virologist opposed to importation, is adamant that he himself could make a definite diagnosis using other tests not requiring live virus, given a little practice. However, even if diagnosis could be made without live virus, Australia's trading partners would need to be convinced that the disease had been eradicated. Laver concedes that this would be difficult without live virus.

The need to use live virus for research, particularly to develop new diagnostic techniques, is also a contentious issue. ASTEC dealt with this need by proposing that a small research group be set up in an overseas laboratory with access to live virus but with ANAH carrying "the main research load". This proposal puts CSIRO in an invidious position. It would be difficult if not impossible to sustain a research programme on this basis, and the overseas group would probably finish up carrying the entire research load.

Meanwhile, the maximum containment facilities at ANAH will be used for research and diagnosis of other exotic viruses and as an investment against a foot and mouth outbreak, when it would be required to handle numerous samples for diagnosis. **Vimala Sarma**

Costly knowledge

YUGOSLAV scientists seem to have been "particularly badly hit" by foreign currency restrictions, according to the Federal Executive Council's Coordination Committee for Science, Technology and Education. The present policy of retrenchment officially called "economic consolidation" has in particular meant the cancellation of virtually all subscriptions to foreign periodicals. Now, the committee says, it has become clear that this consequence is at odds with the government's plan to build up Yugoslavia's own science base and to reduce the country's dependence on imported technology. But since it is cheaper to import journals than to remain dependent on foreign know-how, the coordinating committee has appealed to the government to allot special hard currency funds for subscriptions to science journals and has also begun work on a catalogue of journals in descending order of priority.

Vera Rich

Correction

In the account of work on *in vitro* fertilization in Australia (12 May, p.103), the cost per couple should have read A\$2,500, not A\$42,500. □

Soviet applied research

Broader view demanded

SOVIET industrial research institutes are wasting time on minor improvements to existing technologies at the expense of ideas that could lead to fundamental innovations, a *Pravda* editorial alleged last week. In many cases, the newspaper said, the institutes cannot resist the "commercial contract element" and are inundated by a flood of minor projects. As a result, the proportion of genuine research work done in engineering-oriented institutes has decreased over the past decade by a factor of three.

During the Brezhnev era, it was standard practice to attribute Soviet technological lags to delays in implementing research results. At the time, this explanation seemed entirely reasonable, since the Soviet system of production targets runs into difficulties if innovation involves major retooling; factory managers are not keen on shutdowns which bring production deficits and the loss of bonuses.

Although, during the past fifteen years, there have been several attempts to involve scientists more closely in the industrial process, these have aimed at forming closer links between scientific institutes and local industry; often, existing institutes were combined into a regional "science centre" so as to provide a research base for local production. Direct contracts between industry or agriculture and such institutes or centres were advocated as an effective method of speeding applied research.

Now a reverse swing is under way. While spending time, money and laboratory space on local contracts, the institutes have fallen behind in their major research commitments. During the past two years, *Pravda* says, only 82 industrial pilot plants were completed out of a planned 127, while the Ministry of Agriculture and the State Supply Service (Gossnab) failed to complete a single project on time. Six projects of the Ministry of Petroleum Production and five each of the Ministry of Fisheries and the Ministry of Chemical Machinery due for completion this year are now irretrievably delayed. And in 40 per cent of cases, the delays were due to waste of efforts on "minor research".

The article, with its call for "ministries, departments and party committees" to "strengthen their leadership" over research organizations, follows the usual pattern of a new policy drive, which will almost certainly result in a considerable cutback of minor contracts. The purpose of the turnabout, however, seems to be not so much of increasing central control over research decision-making, but of inculcating among the heads of research teams a greater awareness of the topicality and scientific value of proposed investigations.

Vera Rich

Biotechnology

Cash pile at Biogen

AMONG biotechnology companies, Biogen NV is the most like a bank. In its annual report for 1982, Biogen says that its financial assets amounted to \$57.5 million, much more than its long-term debt of \$5.8 million at the end of the year. Since then it has raised a further \$54 million by its public sale of stock on 22 March (see *Nature* 14 April, p.565).

Part of the explanation of the company's cash position is that it has been able to finance the construction of its laboratory in Cambridge, Massachusetts, with the help of two low interest bonds arranged by the Massachusetts Industrial Finance Agency in 1981, which provided \$4.5 million. During 1982, interest on the company's financial assets amounted to \$8.5 million, two thirds of what it earned by means of payments for research and development carried out.

Paradoxically, Biogen's cash mountain may explain why the share price has drifted downwards from the issue price of \$23 ten weeks ago. While a large proportion of Biogen's income is investment income, investors may think it preferable to put their money with an institution specialized in investment rather than in biotechnology.

Research and development remains the largest item of expenditure (amounting to £18.4 million in 1982) and spending under this heading will continue to increase for the foreseeable future. The total loss on Biogen's operations last year was £4.7 million (\$0.8 million in 1981) and the company says that annual losses are also an integral part of its plans for the years ahead.

Indeed, the accumulated loss is now almost exactly \$10 million, while Biogen

has also invested more than \$1.5 million in securing patents on its products. These costs are carried in the balance sheet as capital assets whose value is written down each year to allow for the shortening life of the patents. The pattern of Biogen's programme remains concentrated on the exploitation of microbially engineered interferons (which are being tested clinically for use in multiple sclerosis, genital herpes and various forms of cancer, not to mention the common cold), vaccines and blood factors. An agreement to produce interleukin-2 in conjunction with a Japanese company was announced earlier this year.

The spread of the company's research contracts remains somewhat narrow. In 1982, 44 per cent of research income came from the Schering Corporation and 37 per cent from the Japanese company Shionogi Ltd. The annual report emphasizes that research income is inherently unpredictable and that 75 per cent of its income in 1982 came from research projects already completed so that further income (from royalties) will not arise until clinical testing is complete.

Biogen's latest project, announced last week, is a partnership to develop a process for making human tumour necrosis factor in conjunction with Suntory Ltd, the Japanese whisky and pharmaceutical manufacturer. This project will be based on a new laboratory called BIOGENT, at Ghent in Belgium, under the direction of Professor Walter C. Fiers, professor of molecular biology at the University of Ghent and a member of the scientific board of Biogen. □

Nature index of biotechnology stocks

12-Month high	12-Month low	Company	Close previous month	Close 27 May	Change
90½	45¼	A.B. Fortia (Sweden)	76¼	82	+5¾
23¼	16¼	Biogen (USA)	18¼	16¾	-1½
6	3	Bio-Logicals (Canada)	5¾	4¾	-1
13	7¼	Bio-Response (USA)	10 1/8	10¾	+¼
19	11¾	Cetus (USA)	16	16½	+½
14½	9½	Collaborative Research (USA)	10¾	9½	-1½
39¾	15	Damon (USA)	34 7/8	37¾	+3½
41	24½	Enzo-Biochem (USA)	37¼	39¾	+2½
18¾	10¾	Flow General (USA)	12 1/8	13¼	+1 1/8
46½	26 1/8	Genentech (USA)	41	42½	+1½
13½	8¾	Genetic Systems (USA)	10¾	12¾	+2
19½	12¾	Genex (USA)	15¾	18½	+2¾
27½	21¼	Hybritech (USA)	27¼	26¾	-½
22¼	13¼	Molecular Genetics (USA)	16¾	22	+5½
23¼	16	Monoclonal Antibodies (USA)	17½	19¼	+1¾
65¾	42	Novo Industri A/S (Den)	57¼	60¾	+3¾

Closing prices are for the last Friday of the month. For over-the-counter stocks, bid price is quoted; for stocks on the American and New York exchanges, the transaction price. The index stood at 223 on 27 May, compared with 213 a month earlier. Data from E.F. Hutton, Inc.

President Reagan's advice

Keyworth answers the critics

Washington

PRESIDENT Reagan's science adviser, Dr George Keyworth, is credited with having talked a parsimonious administration into giving research and development an unexpected increase of 17 per cent in its 1984 budget proposals. But he has also become a controversial figure, chiding the science community for its reluctance to make hard decisions on priorities and for its lukewarm response to the President's call for help in designing a futuristic defence against nuclear attack. Last week he answered the following questions.

The President's call for scientists to design a defence against nuclear attack has been greeted with some scepticism. Have you sought or found any eminent scientists to head an investigation of the possible technologies?

Keyworth: There has been outspoken criticism from an element of the leadership of American science but I think only a single and perhaps even narrow spectrum is represented. Most has come from those individuals who have also criticized basing modes for the MX missile and the strategic modernization programme, and have been deeply involved in arms control measures for years and years. Many are individuals who decades ago embraced a philosophy of unilateral disarmament.

What about people like Richard Garwin?

Keyworth: Richard Garwin has been far more noted for his criticism of trends we have been taking than for his constructive criticism, certainly in recent years.

He played a crucial role in the development of US strategic weapons and the development of MIRVs.

Keyworth: I can only speak of present years, not history. By the same token some of the most outspoken advocates of unilateral disarmament were the people who helped build the atomic bomb in the first place. The responses we are hearing — not in the media — show that a majority of the technical community is quite supportive. We have been able without any difficulty at all to persuade some very distinguished members of the technical community to help enthusiastically to develop the programme.

Critics ask whether, given the number of warheads possessed by the United States and the Soviet Union and their destructive power, a defence that was not pretty near perfect would be worthless.

Keyworth: A common criticism we hear is that, traditionally, defensive measures are inflexible and lethargic while offensive systems are flexible and difficult to counter. The Maginot Line faced the most flexible offensive system there is: men. Here we have quite the opposite. The ballistic missile is an extremely inflexible, rigid, constrained weapon. When it is fired it follows a rigid trajectory. Even when

manoeuvring re-entry vehicles are used in the future you are still talking about a very tight cone. The vehicle's position in space and time has very few uncertainties after launch. To state categorically that it cannot be done, or that it requires a perfect system, or that a leakage rate of 10 to 20 per cent would still cause devastating damage, is a premature and superficial examination of the problem.

What we want to do is build a system effective enough so that the Soviet Union can no longer practically conceive of an attack on the United States or an ally. I for one believe that that is quite feasible.

Would you agree with the view of many particle physicists that Brookhaven's ISABELLE design is going to be obsolete before it can be built? Ought the United States now commit itself to a more ambitious design such as the 'desertron'?

Keyworth: I tend to lean in the same direction but hesitate to use the word obsolete. It is not so much the design of the machine but the operating parameters that are in question. The ISABELLE project was conceived almost a generation ago in high energy physics, in the mid-1970s. We must look now for the proper step for the United States and for the field as a whole in the context of what are the questions to be asked and what is the state of the theory. Do we want to take a bold new step to maintain the vigour of the field or do we want to focus on keeping the maximum number of people busy? I have no doubt that some very exciting forefront research would come out of ISABELLE. But the community seems to believe that a large increase of energy over anything available today or projected for the future offers by far the greatest excitement, and the possibility of clarifying our understanding of elementary particles.

The Nuclear Science Advisory Committee has just recommended the South-East Universities Research Association (SURA) proposal for an electronic accelerator over that of Argonne. The High Energy Physics Advisory Panel may soon endorse ISABELLE. Will you regard these decisions as final?

Keyworth: Ultimately, the federal government will make the final decision although when the government asks a body of scientists for a recommendation the action it eventually takes is traditionally consistent with the recommendations.

The electron machine issue is indicative of a large concern many of us have about American science. Argonne's concern is largely an institutional one. They are asking whether we should begin another laboratory or put the machine at Argonne in spite of the relative merits of the two proposals. If we are going to improve excellence and expand opportunities in

American science, the single greatest threat to progress is parochial institutional concerns. The USRA proposal was excellent. The accusation that it would entail building a new national laboratory is deceptive because there is a clear distinction between multidisciplinary national laboratories and single discipline facilities such as Fermilab. What we are seeing here is institutional interests versus scientific excellence.

So you take a dim view of Argonne's decision and that of local politicians to take the argument to Capitol Hill?

Keyworth: I have every respect for their desire to compete as effectively and aggressively as they can. But it is most unfortunate for science that they wish to bring the issue into the political rather than the scientific realm.

The president of Catholic University has said that the National Centre for Advanced Materials (NCAM) at Berkeley had a dose of its own medicine when its funds were deleted by Congress and Catholic's own vitreous state laboratory got into the budget without peer review. Is that a bad precedent?

Keyworth: I do not think it was by any means the best way of developing science in America. But let's distinguish clearly between NCAM and the Catholic University effort. One is solely an institutional initiative serving that institution's goals and money. NCAM, on the other hand, was developed as a means of enhancing the nation's position. At the forefront of our science policy is a desire to move our immense capability in science closer to the national needs we face in an era of intense competition.

The purpose of NCAM was to develop a mechanism by which a national laboratory could unify the academic world of materials science and the needs of industry. To say that because NCAM did not come through peer review and was part of administration policy it can be compared with the efforts of a hired lobbyist and a single institution's narrow concerns is fatuous and self-serving.

Why was there such a strong reaction against the NCAM proposal?

Keyworth: The scientific community has become overly entrenched in the concept of equitable distribution. The materials science community doubted that NCAM was the optimal step for the health of materials science, and I would not disagree with that. But the issue is what is best for America *in toto*. Where you allocate federal resources is not a scientific issue but a matter of national priorities.

By the way, the NCAM issue is by no means over. The community is a bit unrealistic in assuming that the money originally set aside for NCAM can somehow be used in a way that suits materials science's own priorities. The money will not be available until we can find another mechanism which meets the same objective.

Peter David

Scientist at the helm

SIR — I would like to clarify some points concerning the recent reorganization at the Competitive Research Grants Office (CRGO) of the US Department of Agriculture (*Nature* 12 May, p.104). I am serving as programme manager for a CRGO programme and have had contacts with the USDA officials mentioned in your article. Although it was not clearly identified, the director's position referred to in the article is for the administrator of a newly created Office of Grants and Program Systems (OGPS), of which CRGO is a part, and therefore is not equivalent to the position of chief of CRGO. I have been assured by Assistant Secretary of Agriculture, Orville Bentley, Deputy Assistant Secretary E.L. Kendrick, and Acting Director of OGPS, Clarence Grogan that an active scientist will be appointed as chief to provide the scientific leadership in CRGO. This chief scientist will be recruited from the academic research community, in much the same manner as the past chiefs of CRGO.

I feel that your article was not entirely accurate in describing the present situation at CRGO and that it served only to create a negative image of CRGO. My impression is that Dr Bentley and his staff are committed to the cause of CRGO and I would urge members of the scientific community to give their support to that effort.

ROBERT B. GOLDBERG

Department of Biology,
University of California, Los Angeles, USA

Soviet exchanges

SIR — Your article of 12 May (p.106) referred to Academician Skryabin's planned discussion with the Royal Society. He did indeed come to see me and we briefly reviewed the operation of the exchange agreement between the Soviet Academy of Sciences and the Royal Society. Your article referred to "political" influences that had two years ago reduced the quota (measured in man-months per year). In fact the quota of 49 man-months in each direction has remained unchanged since the present agreement was signed in 1977. The use of the exchange quota within our agreement over the past three years (April to March) has been as follows:

Man-months (no. of visits in brackets)

	1980-81	1981-82	1982-83
From Soviet Union	18 (14)	44 (24)	29½ (18)
From UK	16½ (13)	19¾ (22)	22¾ (22)

Thus although the quota has not changed, it has been under-used in each of the past three years.

R.W.J. KEAY
(Executive Secretary)

The Royal Society,
London SW1, UK

Data banking

SIR — In response to your news article "Who's banker?" (*Nature* 7 April, p.469), I would like to clarify some statements made about the National Biomedical Research Foundation (of which I am president) and its protein and nucleic acid sequence databases. The protein sequence collection has been funded for many years by the National Institutes of Health and that continuing funding will support preparation of the manuscript for Volume 6 of the *Atlas of Protein Sequence and Structure*. Although the tragic and sudden death in February of Dr Margaret Dayhoff is a great loss to us and to the scientific community, the staff that she trained are extremely competent to continue all aspects of the project, some having had over 14 years experience with us. We intend that the work that Dr Dayhoff inspired and led will continue and expand in the future.

The NBRF nucleic acid sequence database will continue as a special purpose collection in that we only include those sequences that code for proteins. We will continue to make our collection available to subscribers.

Neither the Los Alamos group nor Bolt, Beranek and Newman are primarily interested in proteins and, consequently, the format of the GenBank data is such that computer extraction and translation of the coding regions for proteins are very difficult. Furthermore, we find that GenBank annotations are inconsistent and that coding region sequences have not been verified to ensure that the proteins will translate correctly. The format of our Nucleic Acid Sequence Database makes extraction of protein sequences extremely convenient. All of the protein coding regions are translated and compared with the experimentally determined protein sequences (if known) and with the authors' translations. Thus, our data for protein coding regions are triple checked. Researchers primarily interested in protein coding regions will find that our Nucleic Acid Sequence Database is the most accurate available.

R.S. LEDLEY
(President)

National Biomedical Research
Foundation,
Washington DC, USA

Culture collectors

SIR — The paper by Hughes and Datta, "Conjugative plasmids in bacteria of the 'pre-antibiotic' era" (*Nature* 21 April, p.725) is a good example of the usefulness of public depositories or national culture collections.

Too often collections such as Murray's are lost on the retirement or death of the collector when funds or space are not available to preserve the material.

In order to be useful such collections

need to have good documentation and adequate preservation technique. If the Murray collection had been freeze-dried the survival would probably have been 98 per cent instead of the 70 per cent reported.

On documentation, scientists at Fort Detrick recently found that a heat-sensitive plasmid was responsible for the toxigenic properties of the anthrax bacillus. In subsequent examination of a large number of collection strains only one was found that was devoid of the plasmid. This strain we believe came from Pasteur's laboratory via the Kral collection in Prague to the US Department of Agriculture. It would be nice to be able to say this was Pasteur's vaccination strain produced by "pasteurization" of the anthrax culture but we cannot since the provenance is too sketchy to be sure.

To assist in preserving such scientific heritage, the World Federation of Culture Collections has established a standing committee on preservation of endangered collections. Its chairman is Professor Rita Colwell, Department of Microbiology, University of Maryland, College Park, Maryland 20752, USA.

Any scientist knowing of imminent threats to the existence of well documented collections is urged to communicate with the committee so that timely rescue efforts can be undertaken. Also any scientists developing collections are urged to join the World Federation for Culture Collections in order to share and exchange information and cultures with this worldwide network. A revised directory of the world's culture collections has recently been published and is available from Thomas Roswall, Secretary, UNEP/UNESCO/ICRO Panel on Microbiology, Swedish University of Agricultural Sciences, S-750 07 Uppsala, Sweden.

ROBERT E. STEVENSON

American Type Culture Collection,
Rockville, Maryland, USA

Conscious decision

SIR — I have a comment on John Crook's suggestion in *Nature* of 5 May¹ that in the future some "gadgetry" may be discovered in the human brain which can be associated with consciousness. In my opinion such "gadgetry" has already been found. H.L. Seldon has reported² that the axons in the human auditory cortex often travel in helically-wound bundles. If this is the case in other parts of the human cortex as well, it may be that consciousness is due to quantum mechanical resonance. The modes of consciousness may reflect quantization of angular momentum of total spike energy in such bundles.

NEIL BURTNESS

Sycamore, Illinois, USA

1. Crook, J.H. *Nature* 303, 11 (1983).

2. Seldon, H.L. *Brain Res.* 229, 295 (1981).

Names from all nations

SIR — As one responsible for the United Nations University abstract journal, largely aimed at readers in developing countries, I am not surprised that your readers' interest in Chinese names has now extended to South Indian names (*Nature* 17 February, p.560). As both an editor and UN employee, I can see a need for an international guide to personal names.

For Chinese names, we have had to compile our own guide, but the problem of identifying surnames extends to many more cultures than those of China and South India. The problem is complicated by the fact that surnames are not always those used in formal address.

Having assumed that by surname we mean family name, it is helpful to use the term *formal name* for that name or shortest string of names that can be used following a title such as Dr. For Henry Kissinger and Sun Yat-sen, the surnames are formal names and we may address them formally as Dr Kissinger and Dr Sun. This is important since what we are in fact looking for are formal names, which are usually but not always surnames. Instances where we have encountered this divergence will be identified among the following sample list of customs regarding the use of surnames.

Spanish: Typically three or more names with the last two as surnames, sometimes connected by "y", although the second surname is often dropped or abbreviated. The formal name begins with first surname and would include second surname in very formal usage.

Icelandic: Usually two names but no surname. First is formal name in Iceland but second accepted as formal internationally.

Hungarian: Traditionally two names with surname-formal name first but second accepted as formal internationally.

Arabic: Highly variable, often with many names. Formal name often consists of two or three names including articles which can be joined together.

Burmese: One, two or three names typical, but no surname. All names included in formal name.

Thai: Usually two names with surname last but formal name first.

Vietnamese: Usually two or three names. First is surname-formal name.

Japanese: Two names with surname-formal name first when rendered in Japanese but almost always reversed in non-East Asian languages or scripts.

Korean: Usually three names with surname-formal name first. First is sometimes placed last in non-East Asian situations. Same as Chinese usage.

Indonesian: Great variety due to mosaic of ethnic, religious and class groups. For Javanese sometimes one name for all purposes. For some groups including Sundanese, two names often used with first as formal name. Very complex.

African: Great variety due to mosaic of

ethnic, religious and colonial traditions In Rwanda, typically two names with surname last but formal name first.

This sample is far from exhaustive but should give an idea of the diversity involved. Mention could also be made of the problems of identifying titles which are often attached to names, as in the case of U Thant where U means roughly uncle.

Perhaps an international effort to produce a definitive guide would go a long way towards saving much embarrassment in the future. The very diversity of names reflects the fact that human institutions are also very diverse, reflecting the mental constructions that societies have developed to handle the problem of naming their members. We can only hope that the computer forms M.V. Ramana describes will not force all this diversity into monotonous uniformity.

WALTER SHEARER

United Nations University,
Tokyo, Japan

West is east?

SIR — The Greenwich meridian fairly neatly divides the Antarctic continent into an eastern half and a western half. It is in this manner that the place-names "East Antarctica" and "West Antarctica" have come about. Logical as these names seem at first sight, there is no doubt they do give rise to ambiguities, for example "East Antarctica" is to the westward for Australians and New Zealanders; nor does the meridian follow the geographical dividing line of the Transantarctic Mountains and the east side of the Filchner Ice Shelf. The problem is highlighted by Dennis E. Hayes at the end of an introduction to a report¹ of the Deep Sea Drilling Project of *Glomar Challenger*, 1972–73, when this research vessel was working in the Southern Ocean:

"In parts of East Antarctica West Antarctica is east, in others west. This of course depends on if you are in the east East Antarctica or west. However, if you are in west West Antarctica, East Antarctica is west unless you want to go to west East Antarctica in which case it is east. The same holds for east West Antarctica only in the reverse except that if you want to go to west East Antarctica, you still go east. No wonder we don't know what we found!"

May I therefore draw the attention of your readers and contributors² to the alternatives of "Greater Antarctica" and "Lesser Antarctica"; which in addition to being more fitting from the geographical aspect are also more easily remembered. They have been defined as follows³:

Greater Antarctica: A general name for the major region of Antarctica lying in the section of the Indian Ocean side of the Transantarctic Mountains.

Lesser Antarctica: A general name for the minor region of Antarctica lying on the Pacific Ocean side of the Transantarctic Mountains. Lesser Antarctica includes the Antarctic Peninsula.

These definitions were adopted in 1960 by the UK Antarctic Place-names Committee, by the Antarctic Names and Medal Committee of Australia and by the New Zealand Antarctic Place Names Committee. The US Advisory Committee on Antarctic Names approved the names East Antarctica and West Antarctica in 1962, which it felt to be firmly established among scientists and others. The late Dr Brian Roberts, one of the leading architects of the Antarctic Treaty, believed that with time a consensus would evolve in the light of common usage by the scientists and other workers in this field.

ANN SAVOURS

Old Royal Observatory,
National Maritime Museum,
London SE10, UK

1. Hayes, D.E. *Init. Rep. DSDP* 28, 1 (1975).

2. *Nature* 301, 551–552 (1983).

3. Roberts, B.B. *Br. Antarctic Survey Bull.* 53, 257–260 (1981).

Is anybody there?

SIR — I am pleased that the search for extraterrestrial intelligence (SETI) is beginning again in earnest. The current effort by Harvard, Stanford, Berkeley and the National Aeronautics and Space Administration is exciting. It is logical to search in the microwave region near the hydrogen fine-structure wavelength of 21 cm since the phenomenon leading to this emission will be known by technologically oriented societies.

But I would like to make a complementary SETI proposal. I suggest that a search for extraterrestrial signals be made at extremely low electromagnetic frequencies in the neighbourhood of 10 Hz.

There is a logical basis for this suggestion. The Earth, with its ionosphere, forms a natural detector of electromagnetic radiation. If the Earth-ionospheric system is considered as a conducting sphere surrounded by a conducting shell hundreds of kilometres from the surface of the Earth, it becomes a cavity resonator. One can then calculate the eigen-modes of this system¹. The first few eigenstates occur roughly at 8, 14 and 20 Hz.

This matter is familiar to atmospheric scientists. Most of the signals observed at these frequencies are generated by lightning bolts around the Earth.

A comparison of the microwave and extremely low frequency approaches to SETI points out relative advantages and disadvantages. The microwave scenario leads to highly directional beams which are appropriate if you know where you want to send or receive. On the other hand, extremely low frequency transmissions are by necessity rather isotropic (like a dipole radiator) and can therefore be sent or received fairly universally at a substantial cost in intensity.

FRANK J. LOEFFLER

Department of Physics,
Purdue University,
West Lafayette, Indiana, USA

1. Schumann, W.O. *Z. Naturf.* 7a, 149–154; 250–252 (1952).

Fifty years of benzo(a)pyrene

David H. Phillips*

It is fifty years since the publication of the report on the isolation, from coal tar, and identification of the potent chemical carcinogen benzo(a)pyrene, the culmination of over 10 years of research instigated and directed by E.L. Kennaway. The events leading to that discovery are of interest in themselves. Subsequent progress in unravelling the metabolic fate of polycyclic aromatic hydrocarbons has contributed to our understanding of the mechanism of chemical carcinogenesis.

IN the nineteenth century, high incidences of skin cancer were reported among workers in the paraffin refining¹, shale oil² and coal tar industries³. Early attempts to produce cancer in experimental animals with the raw materials of these industries were unsuccessful and it was not until 1915 that the pathologist Yamagiwa and his assistant Ichikawa, at the Imperial University of Tokyo, succeeded in inducing tumours by repeatedly painting the ears of rabbits with coal tar⁴. Soon afterwards Tsutsui⁵ obtained malignant skin tumours in mice by painting them with coal tar. In both cases application of the tar every 2–3 days for several months was required.

E.L. Kennaway joined the Research Institute of the Cancer Hospital (Free) (subsequently the Chester Beatty Research Institute, now the Institute of Cancer Research) in London in 1922 and began his attempts to characterize the carcinogen in coal tar. All that was known then was that the carcinogenic factor was concentrated in the high-boiling fractions and was free of arsenic, nitrogen and sulphur, as had been shown by Bloch and Dreyfuss⁶ in Zurich in 1921. Kennaway, with his assistant F. Goulden, attempted to make tars that were carcinogenic when painted on mouse skin, by heating a variety of materials to high temperature under hydrogen. Positive results were obtained from experiments with isoprene, acetylene, cholesterol, yeast, human hair, muscle and skin⁷. He knew from the classical paper of Berthelot⁸ that acetylene condenses to polycyclic aromatic compounds in such conditions and he had noted from the literature that carcinogenic coal tars contained more aromatic compounds and less paraffins than non-carcinogenic ones⁹. He concluded, therefore, that the carcinogenic agent was some complex polycyclic aromatic hydrocarbon (PAH). He was also able to obtain carcinogenic mixtures by treating tetralin with aluminium chloride at 30–40°C, a process that Schroeter¹⁰ had shown to yield complex aromatic compounds.

In 1924, I. Hieger joined Kennaway and collaborated on studying the Schroeter reaction. They were joined in 1926 by the

physicist W.V. Mayneord who began to study the intense fluorescence noted in the Schroeter fractions. Mayneord made the important observation that the same distinct bands were present in the fluorescence spectra of both the carcinogenic Schroeter fractions and Kennaway's synthetic tars, but attempts to match the fluorescence spectrum with any of the known hydrocarbons in coal tar were unsuccessful.

In 1928, having obtained some samples of PAHs from J.W. Cook, an organic chemist at Sir John Cass College (now the City of London Polytechnic), Hieger and Mayneord found that the fluorescence spectrum of one of them, benz(a)anthracene (Fig. 1), was very similar to those of the carcinogenic tars, although shifted to longer wavelengths¹¹, suggesting that the carcinogen contained the benz(a)anthracene nucleus with some additional substituents. Kennaway and Goulden in 1929 then made dibenz(a,h)anthracene, the synthesis of which had been published by Clar¹² in Dresden, and its 3-methyl derivative, as described by Fieser and Dietz¹³ from Bryn Mawr, Pennsylvania. The fluorescence spectra of the two hydrocarbons showed the same characteristic bands as benz(a)anthracene and the tars, this time at intermediate wavelengths. When, in 1930, they were tested by painting on the skin of mice, the compounds were found to produce tumours¹⁴. This was the first demonstration of the carcinogenic activity of pure chemical compounds.

Cook had joined Kennaway's group in 1929 to synthesize PAHs. In a year he had made 60 new compounds and the group had 146 different mouse skin-painting experiments in progress¹⁵. Early in 1930, work was begun on the isolation of the coal tar carcinogen. Two tons of pitch were distilled at the Becton works of the Gas Light and Coke Company, and the distillate extracted with alcohol. Hieger then successively purified the waxy extract by fractional distillation, differential extraction and crystallization. Fractions were analysed for carcinogenic activity on mouse skin and for their fluorescence spectra; the latter test was "the single thread that led all through the labyrinth"¹⁵ as it greatly simplified the task of identifying likely carcinogenic fractions. In the autumn of 1931

about 7 g of a yellow crystalline material of melting point 116°C had been isolated that was highly carcinogenic and exhibited the correct fluorescence.

C.L. Hewett then joined the group and purified another batch of this material further, obtaining crystals which melted at 160°C, from which Cook isolated two pure products (melting points 176°C and 187°C) and found them to be isomeric with the pentacyclic aromatic hydrocarbon, perylene (C₂₀H₁₂). Cook and Hewett synthesized for comparison the then unknown compounds benzo(a)pyrene (BP) (melting point 177°C) and benzo(e)pyrene (melting point 187°C) and found them to be identical with the major and minor components, respectively, of the mixed crystals. Moreover, both the synthetic and the isolated samples of BP were highly carcinogenic. The results of these experiments were published in April 1933 by Cook, Hewett and Hieger¹⁶, and although Kennaway was not an author of this paper, it is clear that the initiative behind the work was entirely his, despite his having suffered from the debilitating symptoms of Parkinson's disease since 1929 (ref. 17).

Kennaway's contribution and that of Mayneord were acknowledged by the award in 1939 of the first Anna Fuller Memorial Prize jointly to Cook, Hewett, Hieger, Kennaway and Mayneord, "in recognition of their notable accomplishments in the fields of cancer research, specifically for the isolation and synthesis of cancer-producing hydrocarbons from

1. Volkmann, R. *Beiträge zur Chirurgie* (Breitkopf und Hartel, Leipzig, 1875).
2. Bell, J. *Edinb. med. J.* **22**, 135–137 (1876).
3. Butlin, H.T. *Br. med. J.* **11**, 66–71 (1892).
4. Yamagiwa, K. & Ichikawa, K. *Verh. Jap. path. Ges.* **5**, 142–148 (1915).
5. Tsutsui, H. *Gann* **12**, 17–21 (1918).
6. Bloch, B. & Dreyfuss, W. *Schweiz. med. Wschr.* **51**, 1033–1037 (1921).
7. Kennaway, E.L. *J. Path. Bact.* **27**, 233–238 (1924).
8. Berthelot, M. *C. r. hebdom. Séanc. Acad. Sci., Paris* **63**, 479–508 (1866).
9. Kennaway, E.L. *Br. med. J.* **i**, 564–567 (1924).
10. Schroeter, G. *Brennst.-Chem.* **1**, 39–41 (1920).
11. Hieger, I. *Biochem. J.* **24**, 505–511 (1930).
12. Clar, E. *Ber. dt. chem. Ges.* **62**, 350–359 (1929).
13. Fieser, L.F. & Dietz, E.M. *Ber. dt. chem. Ges.* **62**, 1827–1833 (1929).
14. Kennaway, E.L. & Hieger, I. *Br. med. J.* **11**, 1044–1046 (1930).
15. Kennaway, E. *Br. med. J.* **11**, 749–752 (1955).
16. Cook, J.W., Hewett, C.L. & Hieger, I. *J. chem. Soc.* 395–405 (1933).
17. Cook, J.W. *Biogr. Mem. Fellows R. Soc.* **4**, 139–154 (1958).

*Chester Beatty Laboratories, Institute of Cancer Research: Royal Cancer Hospital, Fulham Road, London SW3 6JB, UK

coal tar, their identification by fluorescence spectroscopy, and for the study of the biological effects of these substances".

Structure and activity

In the decade after the isolation of BP, several hundred PAHs were synthesized and tested for carcinogenicity in mice by topical application or by subcutaneous injection. The work of Kennaway's group during this period is summarized in six papers¹⁸⁻²³. While benz(a)anthracene itself had only very weak activity, substitution by methyl groups at some positions produced potent carcinogens, a phenomenon for which there is still no satisfactory explanation. Thus, Bachmann (University of Michigan) synthesized 7,12-dimethylbenz(a)anthracene and 7,8,12-trimethylbenz(a)anthracene²⁴; both were highly active in inducing tumours in mouse skin²⁵ and have since been used widely for the induction of mammary tumours in rats using a single oral dose²⁶. Another potent carcinogen, 3-methylcholanthrene, was obtained by a series of reactions from a bile acid, deoxycholic acid^{20,27}; this aroused interest in the possible endogenous formation of carcinogens of this class and a still unsubstantiated explanation for the occurrence of 'spontaneous' tumours.

The fervour for investigating the properties of PAHs even led Cottini and Mazzone²⁸ to paint the skin of human subjects with BP. The justification for this dubious experiment was the observation made by Haddow²⁹ in 1935 that injections of BP into experimental animals caused a decrease in the growth rate or an actual regression of transplanted tumours. Accordingly, 26 patients with various skin complaints were given daily treatments of BP for up to 4 months. BP caused pigmentation and veruccae in normal skin, which regressed after cessation of treatment, but it was concluded that BP would be carcinogenic if applied to human skin for more protracted periods. There was little therapeutic effect and in some cases the symptoms were aggravated. It is unlikely that such a trial would be sanctioned now.

Several attempts were made to correlate chemical structure with biological activity.

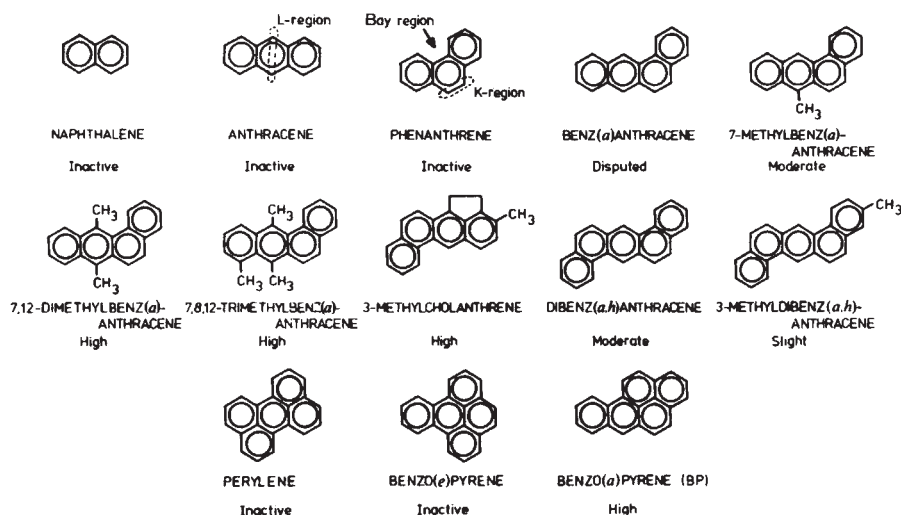


Fig. 1 Structural formulae and carcinogenic activity of PAHs mentioned in the text. Most of the compounds were tested in mice by subcutaneous injection, giving rise to sarcomas, and by topical application to the skin, giving rise to papillomas and epitheliomas. Classification of activity is that of Dipple¹³⁹ and is based on the percentage of treated animals that developed tumours: up to 33%, slight; 33-66%, moderate; > 66%, high. Although inactive as complete carcinogens, phenanthrene, benz(a)anthracene and benzo(e)pyrene have tumour-initiating activity¹⁴⁰.

It was noted early on that most carcinogenic PAHs contain the angular three-ringed phenanthrene unit. The bond equivalent to the 9,10 bond in phenanthrene, called the K-region (Fig. 1) by Pullman³⁰, was identified as possessing greater "double-bond character" and chemical reactivity than other parts of such molecules³¹. Schmidt³² proposed an early qualitative model to explain the significance of this region, that PAH structures contained units of special stability and that, in general, the K-region could not be included in any such unit, but this theory took no account of the delocalized π -electrons which characterize aromatic systems. It was Svartholm³³ in 1941 who first used quantum mechanical methods to infer the electronic properties of the molecules. The French theoretical chemists A. and B. Pullman and R. and P. Daudel, developed Svartholm's model and made it quantitative (reviewed by Coulson³⁴ and the Pullmans³⁵). Although the calculations appeared at one stage to correlate carcinogenic activity with high electron den-

sity at the K-region, as anomalies emerged it became necessary to incorporate the concept of an unreactive L-region (equivalent to the meso 9- and 10-positions of anthracene, Fig. 1) which, in combination with a reactive K-region, would result in carcinogenic activity³⁵. But as time went on, more and more compounds were tested that did not display their predicted carcinogenicity and the theory became untenable. It is now clear that it was based on the false premise that the parent PAHs themselves were biologically active and that their metabolism was a purely detoxifying process.

As early as 1877, it had been found that the urine of naphthalene-fed animals liberated naphthalene on boiling with acid³⁶. By the mid-1930s several metabolites had been identified as specific derivatives of PAHs³⁷⁻³⁹. Further studies revealed that the main metabolites were *trans*-dihydrodiols and phenols, their glucuronic acid and sulphuric acid conjugates, and mercapturic acids (later shown to be derived enzymatically from glutathione conjugates); metabolism was found to affect many parts of the hydrocarbon ring structures and not just the chemically reactive K-region. In 1950, Boyland⁴⁰ (of the Chester Beatty Research Institute) proposed that all the metabolites could theoretically be formed from epoxides and that such intermediates might be responsible for the biological activity of the hydrocarbons. In the absence of direct evidence for the formation of hydrocarbon epoxides, however, Boyland and others continued to subscribe to the theory that carcinogenesis by PAHs was due to their physical intercalation into DNA⁴¹, an idea that persisted at least until 1968⁴², partly because it was almost impossible to either prove or disprove it. But this mechanism could not account for the carcinogenicity of the many other classes of chemical car-

18. Cook, J.W., Hieger, I., Kennaway, E.L. & Mayneord, W.V. *Proc. R. Soc. B111*, 455-484 (1932).
19. Cook, J.W. *Proc. R. Soc. B111*, 485-495 (1932).
20. Barry, G. *et al. Proc. R. Soc. B117*, 318-351 (1935).
21. Bachmann, W.E. *et al. Proc. R. Soc. B123*, 343-368 (1937).
22. Badger, G.M. *et al. Proc. R. Soc. B129*, 439-467 (1940).
23. Badger, G.M. *et al. Proc. R. Soc. B131*, 170-182 (1942).
24. Bachmann, W.E. & Chermida, J.M. *J. Am. chem. Soc.* **60**, 1023-1026 (1938).
25. Bachmann, W.E., Kennaway, E.L. & Kennaway, N.M. *Yale J. Biol. Med.* **11**, 97-102 (1938).
26. Huggins, C.B. *Experimental Leukemia and Mammary Cancer: Induction, Prevention, Cure* (University of Chicago Press, 1979).
27. Wieland, H. & Dane, E. *Hoppe-Seyler's Z. physiol. Chem.* **219**, 240-244 (1933).
28. Cottini, G.B. & Mazzone, G.B. *Am. J. Cancer* **37**, 186-195 (1939).
29. Haddow, A. *Nature* **136**, 868-869 (1935).
30. Pullman, A. *Bull. Cancer* **33**, 120-130 (1946).
31. Clar, E. *Polycyclic Hydrocarbons* Vols 1, 2 (Academic, London, 1964).
32. Schmidt, O. *Naturwissenschaften* **29**, 146-150 (1941).
33. Svartholm, N.V. *Ark. Kemi Miner. Geol.* **A15**, no. 13 (1941).

34. Coulson, C.A. *Adv. Cancer Res.* **1**, 1-56 (1953).
35. Pullman, A. & Pullman, B. *Adv. Cancer Res.* **3**, 117-169 (1955).
36. Baumann, E. & Herter, E. *Hoppe-Seyler's Z. physiol. Chem.* **1**, 244-269 (1877).
37. Bergell, P. & Pschorr, R. *Hoppe-Seyler's Z. physiol. Chem.* **38**, 16-38 (1903).
38. Bourne, M.C. & Young, L. *Biochem. J.* **28**, 803-808 (1934).
39. Boyland, E. & Levi, A.A. *Biochem. J.* **29**, 2679-2683 (1935).
40. Boyland, E. *Biochem. Soc. Symp.* **5**, 40-54 (1950).
41. Boyland, E. & Green, B. *Br. J. Cancer* **16**, 347-360 (1962).
42. Arcos, J.C. & Argus, M.F. *Adv. Cancer Res.* **11**, 305-471 (1968).
43. Scarle, C.E. (ed.) *Chemical Carcinogens* ACS Monograph Ser. no. 173 (American Chemical Society, 1976).
44. Miller, J.A. & Miller, E.C. in *Jerusalem Symposia on Quantum Chemistry and Biochemistry* Vol. 1 (eds Bergmann, E.D. & Pullman, B.) 237-262 (Israel Academy of Sciences and Humanities, Jerusalem, 1969).
45. Miller, J.A. *Cancer Res.* **30**, 559-576 (1970).
46. Miller, E.C. *Cancer Res.* **11**, 100-108 (1951).
47. Miller, E.C. & Miller, J.A. *Cancer Res.* **7**, 468-480 (1947).

cinogens that had then been discovered⁴³ and which lacked the properties characteristic of intercalating agents.

It was at about this time that the Millers^{44,45} proposed what has become a central tenet of present thinking on the mechanism of cancer induction by chemicals: that chemical carcinogens are, or are converted by metabolism into, electrophilic reactants that exert their biological effects by covalent interaction with cellular macromolecules, the critical target most probably being DNA. The Millers had shown that aromatic amines were activated via *N*-hydroxy metabolites (proximate carcinogens) to reactive esters (ultimate carcinogens). A proximate carcinogen might therefore be expected, when metabolized, to exhibit biological activity greater than that of its parent compound, and an ultimate carcinogen to be an electrophile and biologically active as such.

Covalent interactions

In the case of BP (and of PAHs in general), it was some time before the link between metabolism, covalent binding and biological activity was firmly established and the ultimate carcinogen identified.

In 1951, E.C. Miller⁴⁶ (University of Wisconsin) reported that BP, detected by its intense fluorescence, became covalently bound to protein in the skin of mice treated with the hydrocarbon. The observation followed on from her earlier work with J.A. Miller⁴⁷ in which they demonstrated the persistent binding of carcinogenic azo dyes to rat liver protein *in vivo*. In 1961, Heidelberger and Davenport⁴⁸ (University of Wisconsin) demonstrated the covalent binding of dibenz(*a,h*)anthracene to DNA in mouse skin and in 1964 Brookes and Lawley⁴⁹ at the Chester Beatty Research Institute demonstrated a correlation between the carcinogenic potency of six hydrocarbons and the extent to which they became covalently bound in mouse skin to DNA, but not to RNA or protein. That provided important evidence to support the growing conviction that DNA is the critical target for carcinogenesis.

Metabolism studies had progressed, meanwhile, from the use of whole animals to using hepatic microsomal fractions, rich in the mixed-function monooxygenases that metabolize PAHs and other xenobiotics⁵⁰, and Conney *et al.*⁵¹ had shown that treatment of rats with PAHs markedly increases the levels of the enzymes that metabolize them. It was clear by now that the chemical reactivity of the PAHs themselves could not account for the covalent binding observed in animal tissues, but it was not until the late 1960s that direct evidence of the importance of metabolic activation was obtained when Grover and Sims⁵² (Chester Beatty Research Institute) and Gelboin⁵³ (US National Institutes of Health) showed that BP and other PAHs became bound to DNA *in vitro* only in the presence of active metabolizing systems, in these cases pro-

vided by the inclusion of induced rat liver microsomal fractions in the incubation mixture.

Although the metabolic activation of PAHs to DNA-binding products was firmly established, the biological significance of this covalent binding was not. What probably finally convinced the sceptics was the discovery of metabolites having biological activity and the demonstration, in 1973, that many chemical carcinogens are mutagenic for bacteria in the presence of the same metabolizing systems that can catalyse covalent binding to cellular macromolecules. B.N. Ames at Berkeley had developed several histidine-requiring (*His*⁻) strains of *Salmonella typhimurium* which, when treated with a series of chemical carcinogens, including BP and other PAHs, in the presence of rat or human liver homogenates, underwent back-mutation to a *His*⁺ wild type⁵⁴; this is the basis of the Ames test, now widely used to screen chemicals for potential carcinogenic activity. The evident link between metabolism of PAHs and biological activity thus led to the concentration of efforts to identify the metabolites responsible.

K-region epoxides

Although the formation of epoxide intermediates had been proposed in 1950⁴⁰ and the results of the metabolism studies of Boyland, Sims and others were wholly consistent with this hypothesis⁵⁵, conclusive evidence of epoxide formation was not obtained until 1968 when Jerina *et al.*⁵⁶ (NIH) detected the formation of naphthalene 1,2-oxide from naphthalene in a microsomal system by using a radiotracer trapping technique. This was followed by reports⁵⁷⁻⁶⁰ of the formation of epoxides from carcinogenic PAHs, including BP. All the hydrocarbons examined formed epoxides at the K-region of the molecules which, in view of the now obsolete electronic structure correlations³⁵, naturally focused attention on the possibility that K-region epoxides are the ultimately reactive forms *in vivo*.

Syntheses of K-region epoxides had first been achieved in 1964 by Newman and Blum⁶¹ at the Ohio State University. Epoxides were found to react with nucleic acids and protein in the absence of any metabolizing systems⁶², to cause malignant transformation of rodent cells in culture⁶³ and to be mutagens for mammalian cells⁶⁴ and bacteria⁶⁵. These were all properties to be expected of candidate ultimate carcinogens, but tests for carcinogenicity revealed that the K-region epoxides were weaker carcinogens than their parent hydrocarbons⁶⁶⁻⁶⁸. Subsequently, Baird *et al.* found that the chromatographic profiles of hydrocarbon-nucleoside adducts in hydrolysates of DNA from cells treated with PAHs differed from those of hydrolysates of DNA that had been reacted *in vitro* with the K-region epoxide of the hydrocarbon, first for 7-methylbenz(*a*)-anthracene⁶⁹ and then for BP⁷⁰. Thus, the

major metabolites responsible for the covalent modification of DNA and, by inference, for the biological activities of the parent PAHs, were not K-region epoxides. This path of research appeared to have reached a dead end.

Activation via diol-epoxides

Although K-region epoxides seemed not to be responsible for the covalent binding of PAHs to DNA, the involvement of other, as yet unidentified, epoxides was not ruled out. As oxidative metabolism occurs at many sites in PAH, the transient formation of non-K-region epoxides was indeed likely. However, Baird's adduct profiles^{69,70} did not suggest a simple non-K-region epoxide because the hydrocarbon-nucleoside adducts present in digests of DNA from treated cells were more polar than would be expected for those derived from the reaction of a simple epoxide with DNA.

The paper which gave the vital clue to the problem appeared in 1973 from Crocker's laboratory⁷¹ in San Francisco. When BP or various metabolites were incubated with DNA in the presence of hamster liver microsomal fractions, one metabolite, *trans*-BP-7,8-dihydrodiol (Fig.2), became covalently bound to DNA to a 10-fold greater extent than BP, a clear indication that this metabolite is an intermediate in the pathway leading to binding of BP to DNA.

At about the same time, Booth *et al.*⁷² showed that non-K-region dihydrodiols of benz(*a*)anthracene and 7,12-dimethylbenz(*a*)anthracene undergo further metabolism to more polar products, possibly via epoxide intermediates. The oxidation of a hydrocarbon to a non-K-region dihydrodiol can create in the substituted ring an olefinic double bond that might be expected to be a prime site for further metabolism; indeed, Booth and Sims⁷³ obtained a vicinal diol-epoxide, the 8,9-diol-10,11-epoxide of benz(*a*)anthracene, by incubating the *trans*-8,9-dihydrodiol with rat liver microsomal fractions. Sims and his group then prepared the 7,8-diol-9,10-epoxide of BP and showed that the chromatographic profiles of DNA digests obtained from cells treated with BP matched

48. Heidelberger, C. & Davenport, G.R. *Acta Un. int. Cancer*, **17**, 55-63 (1961).

49. Brookes, P. & Lawley, P.D. *Nature* **202**, 781-784 (1964).

50. Holtzman, J.L., Gillette, J.R. & Milne, G.W.A. *J. biol. Chem.* **242**, 4386-4387 (1967).

51. Conney, A.H., Miller, E.C. & Miller, J.A. *J. biol. Chem.* **228**, 753-766 (1957).

52. Grover, P.L. & Sims, P. *Biochem. J.* **110**, 159-160 (1968).

53. Gelboin, H.V. *Cancer Res.* **29**, 1272-1276 (1969).

54. Ames, B.N., Durston, W.E., Yamasaki, E. & Lee, F.D. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2281-2285 (1973).

55. Sims, P. & Grover, P.L. *Adv. Cancer Res.* **20**, 165-274 (1974).

56. Jerina, D.M., Daly, J.W., Witkop, B., Zaltman-Nirenberg, P. & Udenfriend, S. *J. Am. chem. Soc.* **90**, 6525-6527 (1968).

57. Grover, P.L., Hewer, A. & Sims, P. *FEBS Lett.* **18**, 76-80 (1971).

58. Grover, P.L., Hewer, A. & Sims, P. *Biochem. Pharmacol.* **21**, 2713-2726 (1972).

59. Selkirk, J.K., Huberman, E. & Heidelberger, C. *Biochem. biophys. Res. Commun.* **43**, 1010-1016 (1971).

60. Wang, I.Y., Rasmussen, R.E. & Crocker, T.T. *Biochem. biophys. Res. Commun.* **49**, 1142-1149 (1972).

those of digests of DNA reacted with the diol-epoxide. This result, which suggested that *trans*-BP-7,8 dihydrodiol was the proximate carcinogen, and that a 7,8-diol-9,10-epoxide was the ultimate carcinogen of BP, was presented by Sims⁷⁴ at the XIth International Cancer Congress in 1974, and published in *Nature* in the same year⁷⁵.

The next five years saw an explosion of activity in many laboratories aimed at confirming this pathway and elucidating its finer details (Fig.2). The absolute stereochemistry of each active intermediate has been determined and biological activity confirmed in carcinogenicity and mutagenicity (both bacterial and mammalian) tests. Physicochemical studies of the interaction of BP metabolites with DNA *in vivo* and *in vitro* have provided further supportive evidence.

Spectrophotofluorimetry of DNA isolated from tissues or cells treated with BP has shown that the covalently bound BP retains an intact pyrene aromatic nucleic and has thus undergone metabolism in the 7,8,9,10-ring^{76,77}. When radiolabelled samples of *trans*-BP-7,8- and *trans*-BP-9,10-dihydrodiol were applied to mouse skin, only the former gave rise to hydrocarbon-nucleoside adducts, and these were chromatographically indistinguishable from adducts obtained from mice treated with BP⁷⁸. *Trans*-BP-7,8-dihydrodiol is more carcinogenic than BP in mouse skin⁷⁹⁻⁸¹, and more potent in inducing malignant transformation in cultured mammalian cells^{82,83} and mutations in *S. typhimurium* in the presence of a microsomal activating system⁸⁴. The stereoselectivity of the activation pathway is demonstrated by the fact that BP is metabolically converted preferentially to (-)-*trans*-BP-7,8-dihydrodiol^{85,86} via (+)-BP-7,8-oxide⁸⁷, each of which is more active than its enantiomer⁸⁷⁻⁸⁹.

Trans-BP-7,8-dihydrodiol can give rise to two diastereoisomeric vicinal diol-epoxides which differ in the displacement of the epoxide function from the plane of the molecule to the same side as the 7-hydroxyl group (the *syn* configuration) or to the opposite side (the *anti* configuration, Fig.2). The chemical reactivity of

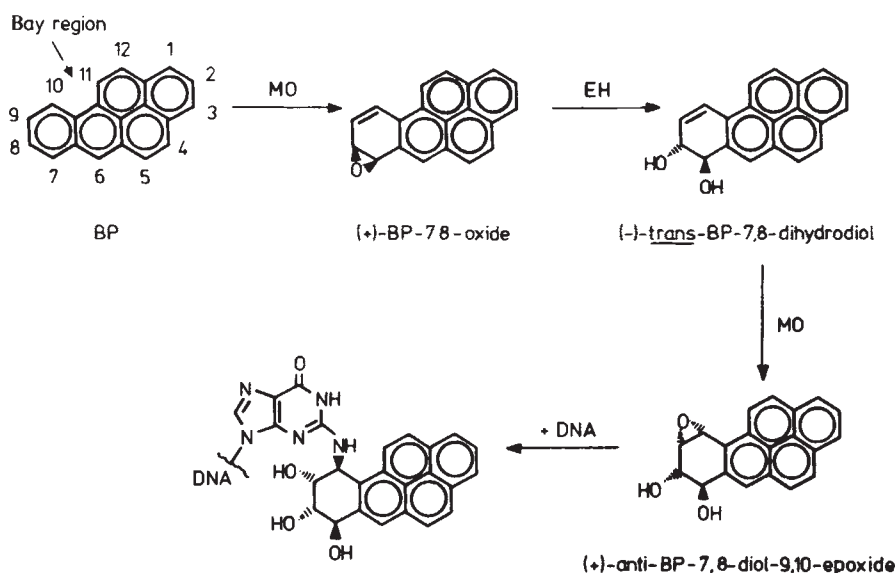


Fig.2 Major pathway of metabolic activation of benzo(a)pyrene. The biotransformations are catalysed by: MO, monooxygenase; and EH, epoxide hydrolase. The structure of the major BP-guanine adduct in DNA is also shown.

these compounds stems from the relative ease with which the strained three-membered epoxy ring will open to generate an electrophilic carbonium ion, and the prediction⁹⁰ that a *syn*-diol-epoxide would form such a carbonium ion more easily than would an *anti*-diol-epoxide because of anchimeric assistance from the favourably positioned 7-hydroxyl group has been confirmed experimentally⁹¹.

Mutagenicity studies on racemic mixtures of the diol-epoxides indicate that the *syn*-isomers are the more potent mutagens to strains of *S. typhimurium*^{92,93}, while the *anti*-isomers are the more active in the mammalian V79 cell system^{92,94-96}. This difference is probably due to the much shorter half life of the *syn*-isomers in aqueous solution⁹⁷, and to their consequent lower concentrations in the mammalian cell nucleus.

There are also significant differences in the mutagenicities and carcinogenicities of the enantiomers of the *anti*- and *syn*-diol-epoxides. Thus, the (+)-*anti*-isomer, the predominant isomer formed metaboli-

cally⁸⁵, and the (-)-*syn*-isomer are more mutagenic than their respective enantiomers⁹⁸. Only the (+)-*anti*-diol-epoxide produced significantly more pulmonary tumours than BP in newborn mice when administered intraperitoneally⁹⁹. Perhaps because of their high chemical reactivity, all the isomers show less activity than BP when applied topically to mouse skin, but again the (+)-*anti*-isomer is the most active^{100,101}.

Physicochemical studies on the interaction of the BP diol-epoxides with DNA and polyribonucleotides have established that they react principally with guanine residues^{102,103}, the C-10 position of BP becoming linked to the exocyclic 2-amino group of the base¹⁰⁴⁻¹⁰⁸ (Fig.2). Adducts formed in tissues or cells treated with BP are derived predominantly from the (+)-*anti*-diol-epoxide, with minor involvement of the (+)-*syn*-diol-epoxide in some cases^{104, 109-112}.

The exact nature of the major BP-DNA interaction and how it may cause a mutagenic event, or initiate a carcinogenic one, is still a matter for speculation. It is even possible that a minor product of the reaction between BP diol-epoxides and DNA^{102, 112-116} is responsible — assuming that DNA is the critical target. Electrical linear dichroism studies¹¹⁷ suggest that the angle between the long axis of DNA and the plane of the bound diol-epoxide is 35° and that the bound carcinogen therefore lies in the minor groove of DNA. Theoretical studies suggest that such a conformation would cause minimal distortion of the DNA double helix¹¹⁸.

Another model proposes that the guanine-deoxyribose bond is rotated through 180°, placing the BP moiety in the major groove, again with little distortion of the DNA structure¹¹⁹. Inversion of the guanine base could conceivably lead to a transversion mutation by mispairing with

61. Newman, M.S. & Blum, S. *J. Am. chem. Soc.* **86**, 5598-5600 (1964).
62. Grover, P.L. & Sims, P. *Biochem. Pharmacol.* **19**, 2251-2259 (1970).
63. Grover, P.L. *et al. Proc. natn. Acad. Sci. U.S.A.* **68**, 1098-1101 (1971).
64. Huberman, E., Aspiras, L., Heidelberger, C., Grover, P.L. & Sims, P. *Proc. natn. Acad. Sci. U.S.A.* **68**, 3195-3199 (1971).
65. Ames, B.N., Sims, P. & Grover, P.L. *Science* **176**, 47-49 (1972).
66. Miller, E.C. & Miller, J.A. *Proc. Soc. exp. Biol. Med.* **124**, 915-919 (1967).
67. Sims, P. *Int. J. Cancer* **2**, 505-508 (1967).
68. Van Durren, B.L., Langseth, L., Goldschmidt, B.M. & Orris, L. *J. natn. Cancer Inst.* **39**, 1217-1228 (1967).
69. Baird, W.M., Dipple, A., Grover, P.L., Sims, P. & Brookes, P. *Cancer Res.* **33**, 2386-2392 (1973).
70. Baird, W.M., Harvey, R.G. & Brookes, P. *Cancer Res.* **35**, 54-57 (1975).
71. Borgen, A. *et al. J. med. Chem.* **16**, 502-506 (1973).
72. Booth, J., Keysell, G.R. & Sims, P. *Biochem. Pharmacol.* **22**, 1781-1791 (1973).
73. Booth, J. & Sims, P. *FEBS Lett.* **47**, 30-33 (1974).
74. Sims, P. *Proc. 11th int. Cancer Congr.*, Florence, 1974

- Vol.2, 9-13 (1975).
75. Sims, P., Grover, P.L., Swaisland, A., Pal, K. & Hewer, A. *Nature* **252**, 326-328 (1974).
76. Daudel, P., Duquesne, M., Vigny, P., Grover, P.L. & Sims, P. *FEBS Lett.* **57**, 250-253 (1975).
77. Ivanovic, V., Geacintov, N.E. & Weinstein, I.B. *Biochem. biophys. Res. Commun.* **70**, 1172-1179 (1976).
78. Grover, P.L., Hewer, A., Pal, K. & Sims, P. *Int. J. Cancer* **18**, 1-6 (1976).
79. Chouroulinkov, I., Gentil, A., Grover, P.L. & Sims, P. *Br. J. Cancer* **34**, 523-532 (1976).
80. Slaga, T.J. *et al. Cancer Lett.* **2**, 115-122 (1976).
81. Levin, W., Wood, A.W., Yagi, H., Jerina, D.M. & Conney, A.H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3867-3871 (1976).
82. Marquardt, H., Grover, P.L. & Sims, P. *Cancer Res.* **36**, 2059-2064 (1976).
83. Mager, R., Huberman, E., Yang, S.K., Gelboin, H.V. & Sachs, L. *Int. J. Cancer* **19**, 814-817 (1977).
84. Malaveille, C., Bartsch, H., Grover, P.L. & Sims, P. *Biochem. biophys. Res. Commun.* **66**, 693-700 (1975).
85. Thakker, D.R. *et al. Chemico-Biol. Interactions* **16**, 281-300 (1977).
86. Yang, S.K., McCourt, D.W., Leutz, J.C. & Gelboin, H.V. *Science* **196**, 1199-1201 (1977).

another guanine. However, it should be remembered that BP diol-epoxides are also frameshift mutagens and that their modification of superhelical simian virus 40 (SV40) DNA results in conformational changes similar to those seen with intercalating agents¹²⁰, although the expected binding angle for intercalation (close to 90°) has not been observed. It is possible that several binding conformations occur but only some initiate events that result in a permanently altered phenotype.

Predictions of the sites of metabolic activation of other carcinogenic PAHs have been made by Jerina and colleagues^{121,122}. The C-10 position of BP, at which carbonium ion formation occurs, can be envisaged as being adjacent to the 'bay-region' of the molecule, that is, the angle formed between benzo rings fused in a nonlinear arrangement. The simplest molecule with a 'bay region' is phenanthrene, in which it lies between the 4- and 5-positions (Fig.1); in BP it is bordered by the 10- and 11-positions (Fig.2). Quantum mechanical calculations predict that a carbonium ion is more readily formed at benzylic positions of saturated benzo rings that are adjacent to 'bay regions' than at non-bay regions for a given hydrocarbon. Since, as we have seen, ring-saturation and carbonium ion formation are both achieved by diol-epoxide formation, the theory predicts greater reactivity for 'bay region' diol-epoxides than non-bay region diol-epoxides, and this has been found to be the case⁹¹. The theory does not, however, predict the carcinogenic potency of a hydrocarbon, which will depend both on the reactivity of the ultimate carcinogen and how much of it is formed metabolically. Nevertheless, 'bay region' diol-epoxides do seem to have an important role in the metabolic activation of the 20 or more PAHs whose metabolites have been examined for biological activity and DNA binding capability so far^{123,124}.

The metabolic activation of chemical carcinogens can be viewed as an error in the process by which organisms detoxify and excrete foreign compounds. The enzymes responsible efficiently metabolize most compounds to less harmful and more readily excretable derivatives even though the organism and its evolutionary precursors may never have been exposed to them. This very versatility occasionally results in the production of an electrophilic metabolite that is not readily detoxified by further metabolism and that may therefore be able to express its genotoxic potential. Thus, what probably distinguishes the BP-7,8-diol-9,10-epoxides from other electrophilic metabolites of BP, such as BP-4,5-oxide, is that the former are poor substrates for epoxide hydrolase and thus are not as readily detoxified¹²⁵. Other mechanisms of detoxification, such as conjugation with glutathione¹²⁶ and non-enzymatic hydrolysis to tetrols^{97,127}, may still have some influence on the balance between activation and deactivation.

The risk to humans

The nature of research is such that BP will continue to be an important research tool for many disciplines. The vast amount of data that has accumulated on the cellular processing of BP¹²⁸⁻¹³⁰ makes it a valuable standard with which new systems can be characterized and evaluated. Much current research simply uses BP in this capacity: tissues and cells can be rapidly assayed for their ability to metabolize BP; originally by means of the strong fluorescence of one metabolite, 3-hydroxybenzo(a)pyrene¹³¹, more recently by the use of sensitive HPLC techniques to separate and assay complex mixtures of metabolites^{132,133}.

Equally, however, BP is deserving of study because PAHs are now recognized as major environmental pollutants¹³⁴. They are released into the atmosphere during the combustion of fossil fuels and vegetation and are present in tobacco smoke. Exposure to PAHs is virtually unavoidable, and they are strongly suspected of being a causative factor in several human cancers of epithelial origin (for example, skin, lung, bronchus and colon). The metabolism of PAHs, especially BP, in cell and organ cultures of many human tissues^{124,135} gives no evidence to suggest that human tissues would not be susceptible to PAH carcinogenesis: metabolism is qualitatively similar in animal tissues known to be susceptible to

the carcinogens and in the corresponding human tissues, although there are significant quantitative differences between human tissue samples from different people. Whether or not such differences reflect differences in individual human susceptibility to cancer is not yet clear. Some animal studies do not support this concept, however, because inbred strains of mice with different susceptibilities to PAH carcinogenesis nevertheless metabolize PAHs to DNA-binding products to similar extents¹³⁶. Estimating the risk from human exposure to these compounds will probably require more direct methods than tissue culture studies. The production of antibodies to BP-diol-epoxide-DNA adducts¹³⁷ which has made possible the assay of human biopsy material for prior exposure to BP¹³⁸ may — together with prospective epidemiological studies — be a step towards determining the hazard that PAHs present to humans.

Addendum: PAHs have been investigated to a limited extent in the treatment of human cancer. Beginning in 1934, Bauer^{141,142} treated 22 patients with advanced skin cancer with BP and the tumours regressed in seven cases. Also Huggins and McCarthy¹⁴³ treated six cases of metastatic breast cancer with intramuscular injections of 3-methylcholanthrene and observed inhibition of malignancy in five of the patients. □

87. Levin, W. *et al. J. biol. Chem.* **255**, 9067-9074 (1980).
88. Levin, W. *et al. Cancer Res.* **37**, 2721-2725 (1977).
89. Huberman, E., Yang, S.K., McCourt, D.W. & Gelboin, H.V. *Cancer Lett.* **4**, 35-43 (1977).
90. Hulbert, P.B. *Nature* **256**, 146-148 (1975).
91. Yagi, H., Hernandez, O. & Jerina, D.M. *J. Am. chem. Soc.* **97**, 6881-6883 (1975).
92. Wood, A.W. *et al. Cancer Res.* **36**, 3358-3366 (1976).
93. Malaveille, C., Kuroki, T., Sims, P., Grover, P.L. & Bartsch, H. *Mutat. Res.* **44**, 313-326 (1977).
94. Huberman, E., Sachs, L., Yang, S.K. & Gelboin, H.V. *Proc. natn. Acad. Sci. U.S.A.* **73**, 607-611 (1976).
95. Newbold, R.F. & Brookes, P. *Nature* **261**, 52-54 (1976).
96. Marquardt, H., Baker, S., Grover, P.L. & Sims, P. *Cancer Lett.* **3**, 31-36 (1977).
97. Keller, J.W., Heidelberger, C., Beland, F.A. & Harvey, R.G. *J. Am. chem. Soc.* **98**, 8276-8277 (1976).
98. Wood, A.W. *et al. Biochem. biophys. Res. Commun.* **77**, 1389-1396 (1977).
99. Buening, M.K. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 5358-5361 (1978).
100. Levin, W. *et al. Cancer Res.* **37**, 3356-3361 (1977).
101. Slaga, T.J. *et al. Cancer Res.* **37**, 4130-4133 (1977).
102. Jennette, K.W. *et al. Biochemistry* **16**, 932-938 (1977).
103. Meehan, T., Straub, K. & Clavin, M. *Nature* **269**, 725-727 (1977).
104. Weinstein, I.B. *et al. Science* **193**, 592-595 (1976).
105. Jeffrey, A.M. *et al. J. Am. chem. Soc.* **98**, 5714-5715 (1976).
106. Koreeda, M., Moore, P.D., Yagi, H., Yeh, H.J.C. & Jerina, D.M. *J. Am. chem. Soc.* **98**, 6720-6722 (1976).
107. Osborne, M.R., Beland, F.A., Harvey, R.G. & Brookes, P. *Int. J. Cancer* **18**, 362-368 (1976).
108. Nakanishi, K. *et al. J. Am. chem. Soc.* **99**, 258-260 (1977).
109. King, H.W.S., Osborne, M.R., Beland, F.A., Harvey, R.G. & Brookes, P. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2679-2681 (1976).
110. Jeffrey, A.M. *et al. Nature* **269**, 348-350 (1977).
111. Baird, W.M. & Diamond, L. *Biochem. biophys. Res. Commun.* **77**, 162-167 (1977).
112. Koreeda, M. *et al. Science* **199**, 778-781 (1978).
113. Osborne, M.R., Harvey, R.G. & Brookes, P. *Chemico-Biol. Interactions* **20**, 123-130 (1978).
114. Jeffrey, A.M. *et al. Science* **206**, 1309-1311 (1979).
115. Meehan, T. & Straub, K. *Nature* **277**, 410-412 (1979).
116. Osborne, M.R., Jacobs, S., Harvey, R.G. & Brookes, P. *Carcinogenesis* **2**, 553-558 (1981).
117. Geacintov, N.E., Gagliano, A., Ivanovic, V. & Weinstein, I.B. *Biochemistry* **17**, 5256-5262 (1978).
118. Beland, F.A. *Chemico-Biol. Interactions* **22**, 329-339 (1978).
119. Kadlubar, F.F. *Chemico-Biol. Interactions* **31**, 255-263 (1980).
120. Drinkwater, N.R., Miller, J.A., Miller, E.C. & Yang, N.-C. *Cancer Res.* **38**, 3247-3255 (1978).
121. Jerina, D.M. *et al. In Vitro Metabolic Activation in Mutagenesis Testing* (eds de Serres, F.J., Fouts, J.R., Bend, J.R. & Philpot, R.M.) 159-177 (Elsevier, Amsterdam, 1976).
122. Jerina, D.M. & Daly, J.W. in *Drug Metabolism: from Microbe to Man* (eds Parke, D.V. & Smith, R.L.) 13-32 (Taylor and Francis, London, 1977).
123. Sims, P. & Grover, P.L. in *Polycyclic Hydrocarbons and Cancer Vol.3* (eds Gelboin, H.V. & Ts'o, P.O.P.) 117-181 (Academic, New York, 1981).
124. Conney, A.H. *Cancer Res.* **42**, 4875-4917 (1982).
125. Wood, A.W. *et al. J. biol. Chem.* **251**, 4882-4890 (1976).
126. Cooper, C.S., Hewer, A., Ribeiro, O., Grover, P.L. & Sims, P. *Carcinogenesis* **1**, 1075-1080 (1980).
127. Yang, S.K., McCourt, D.W. & Gelboin, H.V. *J. Am. chem. Soc.* **101**, 5130-5134 (1979).
128. Gelboin, H.V. & Ts'o, P.O.P. (eds) *Polycyclic Hydrocarbons and Cancer Vols 1, 2* (Academic, 1978).
129. Gelboin, H.V. *Physiol. Rev.* **60**, 1107-1166 (1980).
130. Cooper, C.S., Grover, P.L. & Sims, P. *Prog. Drug Metab.* **7**, 295-396 (1983).
131. Wattenberg, L.W., Leong, J.L. & Strand, P.J. *Cancer Res.* **22**, 1120-1125 (1962).
132. Selkirk, J.K., Croy, R.G., Roller, P.P. & Gelboin, H.V. *Cancer Res.* **34**, 3474-3480 (1974).
133. Holder, G.M., Yagi, H., Jerina, D.M., Lu, A.Y.H. & Conney, A.H. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4356-4360 (1974).
134. National Academy of Sciences, Committee on Biologic Effects of Atmospheric Pollutants. *Particulate Polycyclic Organic Matter*, Washington, DC (1972).
135. Harris, C.C., Trump, B.F., Graftstrom, R. & Autrup, H. *J. cell. Biochem.* **18**, 285-294 (1982).
136. Phillips, D.H., Grover, P.L. & Sims, P. *Int. J. Cancer* **22**, 487-494 (1978).
137. Poirier, M.C., Santella, R., Weinstein, I.B., Grunberger, D. & Yuspa, S.H. *Cancer Res.* **40**, 412-416 (1980).
138. Perera, F.P. *et al. Carcinogenesis* **3**, 1405-1410 (1982).
139. Dipple, A. in *Chemical Carcinogens* (ed. Searle, C.E.) 245-314 (ACS Monograph Ser. no.173, American Chemical Society, 1976).
140. Scribner, J.D. *J. natn. Cancer Inst.* **50**, 1717-1719 (1973).
141. Bauer, K.H. *Arch. klin. Chir.* **189**, 123-184 (1937).
142. Bauer, K.H., Rarici, B. & Gummel, Arch. klin. Chir. **193**, 499-502 (1938).
143. Huggins, C. & McCarthy, J.D. *Cancer Res.* **17**, 1028-1032 (1957).

What will come after the Z^0 ?

CERN now seems confident that it has observed five examples of the Z^0 particle, a kind of heavy photon. What is there left to do?

IT'S official! CERN, the European centre for nuclear research, has discovered the Z^0 — the particle related to the photon, but heavier, which mediates the "neutral current" component of the weak nuclear force. The Z^0 , with its charged partners the W^+ and W^- , discovered together at CERN earlier this year, was predicted by the Salem-Weinberg model. This model unifies the weak and electromagnetic forces in a single field theory.

A few weeks ago, evidence was found of the production and decay of just one Z^0 , but the evidence was somewhat equivocal and no official announcement was made (see *Nature* 19 May p.193). Now there are five examples, one of the "a text-book version", according to Professor Peter Kalmus of Queen Mary College, London, one of the 130-man, 13-university research team that discovered the particles.

This "UA1" collaboration works with the proton-antiproton collider at CERN, a device which uses the Super Proton Synchrotron (SPS) to accelerate and store protons moving in one direction and antiprotons in the other. Where the beams collide, many kinds of particles are created, among them W s and Z s.

UA1 now has four Z^0 s decaying into an electron and a positron, and one into muon (heavy electron) and antimuon. The mass appears to be "about" 95 GeV — less than the figure rumoured for the first event, and more in line with the Salem-Weinberg model, in relation to the 81 ± 2 GeV quoted in a recent UA1 seminar for the mass of the W^- . The group has roughly 30 examples of production and decay of the W .

Even so, it seems, there will be no published masses until the present data-taking run is complete on 4 July, when the detector will be calibrated. By then the

group would like to have 60–70 W decays and 6–7 Z^0 decays in the bag, with conclusions on masses and fits to the Salem-Weinberg theory by July.

The theory predicts W and Z masses which depend on a number called the "Weinberg angle", which measures quantum mechanical mixing between the raw photon and Z^0 and their physical counterparts. Expressed as $\sin^2 \theta_w$, the presently accepted value is 0.23 ± 0.01 . Using this number, the Salem-Weinberg theory yields a W mass of 82 ± 2 . On the same basis, the theory predicts a Z mass of around 93 GeV, consistent so far with UA1 estimates.

UA1 measurements also show that the W decays are asymmetric, as should be the case because of the weak interaction's violation of the "parity" between right and left. In the W^+ decays found at CERN, for example, the positron in the decay tends to move off in the same hemisphere as the antiproton, as predicted by the theory.

Professor Abdus Salam, speaking from the International Centre for Theoretical Physics in Trieste this week, described himself as "very delighted" with the results. "I have the greatest admiration for the people involved", he said, "and particularly for the construction of the collider".

Salam sees these results in one respect as "the end of the road". Most of the grand unified theories that go beyond the Salem-Weinberg model include colour — the force field of the quarks, which are the components of particles like protons — and are forced to invoke phenomena at extremely high energies (or equivalently tiny distances) to explain the splitting between the fundamental forces we observe. These energies (of 10^{19} GeV or so) are in practice

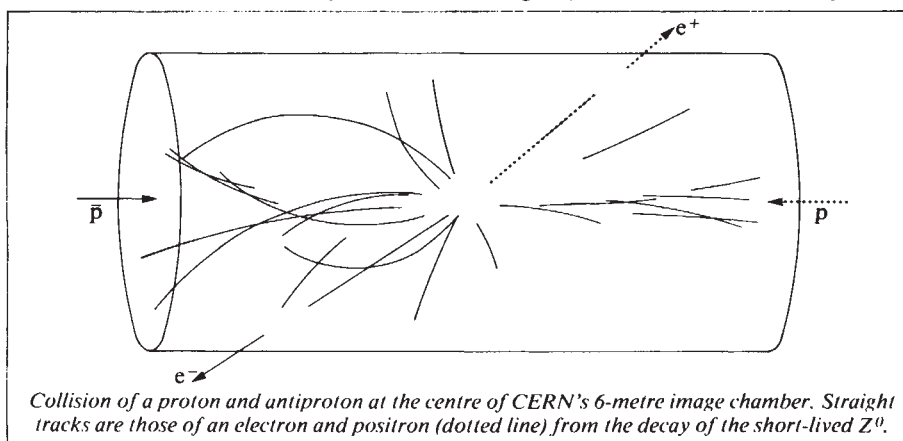
unobtainable with present technologies, so that theories must be tested indirectly, by measurement of the decay of the proton and the like.

However, Salam has another theory, with his collaborator Pati, in which the left-right asymmetry of the weak force is easier to understand... This could be tested with the CERN collider, Salam says. It predicts a second Z^0 , at higher mass (above 200 GeV), which the CERN collider should produce about one time for every fifty times it produces an ordinary Z^0 . The UA1 group could detect it if they could collect enough events — which means increasing the "luminosity" of the colliding beams.

Such an increase is indeed on the cards. According to Kalmus "we had a few events [not including W s and Z s] in 1981; we had 10 times the number in 1982; we might get another factor of 10 this year, and if we get more running time, even another factor of 5–10 the year after". This might be enough to see Salam's second Z^0 . If not, there are plans to improve the production and collection of antiprotons, which is one of the main limitations on the collider luminosity, by building a new collector ring. But this would take three years to build, and take money from CERN's next big project, the large electron-positron collider LEP. This latter is due to produce physics in mid-1989, but since it is to be constructed on a constant (and in real terms slightly reducing) CERN budget, money spent on the collector might delay LEP. Salam's opinion: "I think Rubbia [the collider designer and leader of UA1] should be given every possible help".

According to Kalmus, there is also plenty to do with the collider other than search for Salam's putative second Z^0 : study jets, for example, where a quark emerges from a proton-antiproton collision virtually isolated but then — as quarks cannot exist alone — "dresses itself" in a cloud of other strongly interacting particles, emerging thus as a clump of tracks isolated from others in the events. These jets may hide evidence of heavy quarks (a sixth, called the top quark, is needed to complete the present list of five). They may also conceal "Higgs particles", whose mass is unknown but whose existence is ultimately crucial to the Salem-Weinberg model. Another possibility is to search for Centauro events, seen a few times in cosmic rays, which indicate that a new kind of physics may set in at energies not much above present values.

Robert Walgate



Haemophilia

Benefits of cloning genes for clotting factors

from A.L. Bloom

ABOUT six men in every 100,000 of the population are affected by severe haemophilia. Of these, probably five will suffer from haemophilia A, caused by an X-chromosome-linked deficiency of the procoagulant antihemophilic factor A (factor VIII, VIIIc), and one from the equally severe but less common haemophilia B, which results from an X-linked deficiency of antihemophilic factor B (factor IX). Several papers published over the past year suggest that through the application of the techniques of gene cloning, our understanding of haemophilia should increase considerably and new ways of controlling it may soon be at hand.

Factor IX is synthesized in parenchymal liver cells and the amino acid sequence of the bovine and human protein have been fully established. This knowledge has facilitated the cloning of the gene. Briefly the strategy adopted by Brownlee and colleagues at Oxford¹ involved preparation of bovine liver mRNA enriched for mRNA specific for factor IX. An oligonucleotide, synthesized from knowledge of factor IX's amino acid sequence, was used to prime reverse transcription of the mRNA. The resultant cDNA was used to prepare a plasmid library in *Escherichia coli* and a clone positive for bovine factor IX was detected by means of a second labelled synthetic oligonucleotide probe. The bovine probe was used to screen a cloned genomic human library and a positive clone matching about 25 per cent of the factor IX sequences was isolated. The strategy developed by Kurachi and Davie² involved the transcription of baboon liver polyosomal mRNA immunologically enriched for factor IX. A human genomic DNA library was screened with the labelled baboon cDNA probes and the specificity of positive clones for human factor IX was identified with probes developed from appropriate synthetic oligonucleotides.

The Oxford group has gone on to look for pathological lesions in the factor IX gene of haemophilia B patients, using four different genomic probes³. They wisely reasoned that the patients most likely to have gene deletions would be those who are so devoid of factor IX antigen that they developed antibodies to factor IX after replacement therapy with concentrate. This proved to be the case in four of the five patients studied.

However only about 1 per cent of patients with haemophilia B develop antibodies and further exploratory data will be needed on ordinary patients. Many of these have circulating molecular variants and it

seems likely that these individuals will turn out to have more limited genomic defects. In this case the use of gene probes for carrier detection and prenatal diagnosis may depend, for instance, upon the demonstration of extra- or intragenomic linked polymorphisms in a similar way to that for β -thalassaemia and the haemoglobinopathies⁴.

A second obvious potential benefit from cloning the factor IX gene is the expression of the functionally active molecule in bacteria. Haemophilia B is an uncommon complaint, however, and there is little shortage of factor IX concentrate prepared from blood; so for the moment at least, production of factor IX by engineered bacteria is not commercially viable.

All this brings us to the question of the much more common and hence clinically important haemophilia A. Are parallel developments occurring with factor VIII?

This is not a simple question to answer. Factor VIII is a much more complicated molecule. It has a molecular weight of about 300,000 or more in plasma but may consist of smaller subunits⁵, and is non-covalently linked with an autosomally controlled huge polymeric protein, Willebrand factor (WF, factor VIII-related antigen), so called because it is deficient or abnormal in an autosomally inherited bleeding disorder named von Willebrand's disease. Factor VIII is an extremely labile molecule present only in low concentration in plasma (about 200 ng ml⁻¹), so its isolation and characterization have proved difficult. Furthermore, although it has long been recognized that Willebrand factor is synthesized in endothelial cells of most blood vessels, the specific cellular site of synthesis of factor VIII is quite unknown. Attempts to develop xenogeneic antisera for factor VIII have been singularly unsuccessful and immunohistological studies using human antibodies are fraught with difficulties from non-specific reactions.

The recent development of monoclonal antibodies both to factor VIII (VIIIc)⁶ and to Willebrand factor⁷ may well mark a turning point however. Previous evidence had suggested that factor VIII may be synthesized in the liver⁸; and in this issue of *Nature*, Stel and co-workers⁹ describe the immunohistological localization of factor VIIIc antigen specifically in hepatic sinusoidal endothelial cells. Of course this is a long way from indicating synthesis by these cells, but it is the first reasonably acceptable immunohistological demonstration of the cellular localization of this elusive molecule.

Monoclonal antibodies are also proving

useful for the chemical isolation of factor VIII from plasma. At a recent meeting E.G.D. Tuddenham and colleagues at the Royal Free Hospital in London described the use of preliminary conventional concentration techniques, followed by immunopurification of the whole factor VIII-WF complex on immobilized monoclonal anti-WF (anti-VIIIcAg), ionic dissociation of factor VIII and its final purification on immobilized monoclonal anti-factor VIII antibody.

It is an unfortunate feature of work on factor VIII (and no doubt other proteins) that commercial interests have dictated some degree of secrecy. Thus it is probable that the sequences of at least parts of the human and porcine factor VIII molecules are already available. It is also a fair bet that the factor VIII gene has been or is in the process of being cloned either by strategies similar to those used for factor IX, or by adopting a more empirical approach such as that used successfully for urokinase¹⁰.

The implications of successful cloning of factor VIII are great. In the first place it would provide a diagnostic probe that could be applied in a way analogous to that which is being pursued for factor IX. Thus it could be used to detect gene deletions that may indicate patients at risk from developing antibodies and resistance to treatment. Demonstration of deletions or linked polymorphisms will assist carrier detection and would allow prenatal diagnosis, for example from chorionic biopsy at an early stage of pregnancy¹¹, or even from fetal cells harvested from maternal blood with a fluorescence-activated cell sorter. These procedures would be a distinct advantage over current methods of prenatal diagnosis which, elegant though they are, depend upon technically demanding fetal blood sampling for phenotypic assessment at an inconveniently late stage in the second trimester of pregnancy¹². Labelled cDNA probes for factor VIII will thus provide valuable laboratory reagents.

Expression of haemostatically effective factor VIII in bacteria is another matter, since the structural or post-translational constraints for biological activity have not yet been determined. Separation of the protein from those of the microorganism with the aid of monoclonal antibodies will

1. Choo, K.H., Gould, K.G., Rees, D.J.G. & Brownlee, G.G. *Nature* **299**, 178 (1982).
2. Kurachi, K. & Davie, F.W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6461 (1982).
3. Giannelli, F. *et al.* *Nature* **303**, 181 (1980).
4. Little, P.F.R. *et al.* *Nature* **285**, 144 (1980).
5. Fay, P.J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 7200 (1982).
6. Müller, H.P. *et al.* *Blood* **58**, 1000 (1982).
7. Sola, B. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 183 (1982).
8. Bloom, A.L. *Clinics Haemat.* **8**, 53 (1979).
9. Stel, H.V., Veerman, F.C.L. & van der Kwast, Th. *Nature* **303**, 530 (1983).
10. Ratkin, B. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 3313 (1981).
11. Williamson, R. *et al.* *Lancet* **ii**, 1125 (1981).
12. Mibashan, R.S., Rodeck, C.H. & Thompson, J.K. in *The Hemophilias* (ed. Bloom, A.L.) 176 (Churchill Livingstone, Edinburgh, 1982).

certainly be a help however.

The therapeutic and financial consequences of success in this field will be enormous. Albumin and factor VIII form the stimulus for commercial and national plasma-fractionation programmes. Thus the biogenetic synthesis of either of these proteins will profoundly influence fractionation policy as well as provide

potentially virus-free therapeutic agents. It is not difficult to see why the biotechnical commercial conglomerates jealously guard their secrets to the frustration of mere academics and clinicians. □

A.L. Bloom is Professor in the Department of Haematology, Welsh National School of Medicine, Heath Park, Cardiff CF4 4XN.

Cancer

Chromosome aberrations and oncogenes

from Fred Gilbert

THE status of the chromosome localization of the cellular homologues of retroviral oncogenes (the *c-onc* genes) has recently been summarized in these pages¹. The review points out that the breakpoints of many of the consistent rearrangements that characterize individual cancers involve the specific bands to which different *c-onc* genes have been assigned. Although a role for *c-onc* genes in mammalian cancer has yet to be proved, a possible relationship between chromosome changes, alterations in *onc* gene activity and cellular transformation can be deduced from what is known of the genetics and cytogenetics of cancer in man, and from studies of *onc* genes in man and other species. The observations contributing to the scheme I shall propose can be summarized as follows.

First, chromosome abnormalities in human cancer can be grouped into three major classes: reciprocal translocations, deletions and duplications (the latter including duplications of whole chromosomes, chromosome segments and gene sequences, as in the structural analogues of gene amplification, homogeneously staining regions and double minutes; ref. 2 and F. Gilbert, in preparation). Examples of each of these have now been associated with presumed changes in *onc* gene activity. As well as the translocations and deletions cited by Rowley¹, amplification of an *onc* gene, *c-myc*, has been demonstrated in a human promyelocytic leukaemia cell line, HL-60, known to contain double minutes^{3,4}.

Second, in many, if not all cancers, both tumorigenesis and tumour progression probably involve multiple gene changes. This seems to be true in certain solid tumours of children — retinoblastoma, neuroblastoma and Wilms' tumour — each of which has been associated with chromosome rearrangements involving different specific segments⁵⁻⁷ and for each of which it has been postulated that (at least) two gene changes are required⁸. It is also clear that more than one putative oncogene may be active in a single cancer cell; a number of leukaemias and lymphomas have now been found to contain DNA sequences capable of inducing transforma-

tion of the mouse cell line NIH 3T3 (when transferred by transfection⁹) in addition to the transcriptionally active retroviral *onc* genes previously identified in each^{10,11}.

Third, cellular differentiation can be viewed as a series of step-wise changes in which the potential for (and pattern of) phenotypic expression is altered with successive cell divisions¹². In particular cell lineages, terminal differentiation (the expression of the fully differentiated programme) may also be associated with the cessation of cell division¹³. Control of the individual stages of development is (at least partially) genetic and among the genes which may play a part in this process are the *onc* genes. Recent reports have documented the differential expression of particular *c-onc* genes during pre- and postnatal development in the mouse¹⁴ and during liver regeneration in the rat¹⁵. That there may be a relationship between differentiation, specific chromosome aberrations and malignancy is also suggested by the demonstration that rearrangements involving homologous segments in certain B-cell tumours of man (Burkitt lymphoma) and mouse (plasmacytoma) result in the translocation of the same *onc* gene from one chromosome to another¹.

The synthesis of these observations leads me to propose the following scheme: the development of a chromosome abnormality, if it involves a segment containing or adjacent to a *c-onc* gene, can produce an alteration in the pattern of expression of the gene different from that seen in the normal course of cellular differentiation. The chromosome abnormality may develop as a germinal change¹⁶, a *de novo* event, or following an underlying mutation which increases instability at that site¹⁷. The alteration in *c-onc* gene expression may be manifest in decreased gene activity (in deletions of the structural locus¹⁸), in increased activity (in deletions of a putative *onc* gene suppressor, in duplications, amplification and some translocations¹⁹), or in the transcription of an abnormal mRNA (postulated in other translocations²⁰). Aberrant *onc* gene expression may, in turn, produce an alteration in the pattern of cell division. Instead of the programmed cessa-

tion of division, the cells may continue to divide; with continued division, the statistical likelihood of additional gene changes (through mutations and further chromosome rearrangements) increases.

The accumulation of specific gene changes is postulated to result in tumorigenesis; those changes which occur following tumorigenesis are presumed to contribute to tumour progression. The primary gene changes leading to tumorigenesis are likely to include those resulting from chromosome rearrangements characteristic of particular tumours, such as the 8;14 translocation in Burkitt lymphoma¹. The secondary gene changes involved in tumour progression are likely to include those resulting from chromosome abnormalities that are not unique to any one tumour type: trisomies for chromosomes 17 and 1, for example²¹.

Two caveats must be added: changes in gene activity can clearly be produced by mechanisms other than chromosome rearrangement (for discussions of the possible significance of point mutations in *onc* genes see refs 22,23) and *onc* genes (probably) constitute only one category of genes involved in cancer. However, recombinant DNA probes specific for *onc* genes are available and can be used, in conjunction with permanent cell lines containing defined chromosome rearrangements, to determine precisely how such rearrangements alter the expression of *onc* gene activity in individual cases. The *onc* genes therefore represent a model system with which to study the part played by single-gene changes in the development of cancer. □

Fred Gilbert is Associate Professor of Medical Genetics and Pediatrics, Mount Sinai School of Medicine, New York, New York 10029.

- Rowley, J.D. *Nature* **301**, 290 (1983).
- Balaban-Malenbaum, G. & Gilbert, F. *Science* **193**, 739 (1977).
- Dalla Favera, R., Wong-Staal, F. & Gallo, R.C. *Nature* **299**, 61 (1982).
- Collins, S. & Groudine, M. *Nature* **298**, 679 (1982).
- Yunis, J.J. & Ramsay, N. *Am. J. Dis. Child.* **132**, 161 (1978).
- Gilbert, F. *et al. Cancer Genet. Cytogenet.* **7**, 33 (1982).
- Francke, U., Holmes, L., Atkins, L. & Riccardi, V. *Cytogenet. Cell Genet.* **24**, 185 (1979).
- Knudson, A.G. *Ped. Res.* **10**, 513 (1976).
- Cooper, G. *Science* **217**, 301 (1982).
- Cooper, G. & Neiman, P. *Nature* **292**, 857 (1981).
- Ozanne, B. *et al. Nature* **299**, 744 (1982).
- Holtzer, H. in *Stem Cells and Tissue Homeostasis* (eds Ford, B., Potten, C. & Cole, R.) 1 (Cambridge University Press, 1978).
- Gilbert, F. *J. natn. Cancer Inst.* **57**, 667 (1976).
- Muller, R. *et al. Nature* **299**, 640 (1982).
- Goyette, M., Petropoulos, C., Shank, P. & Fousto, N. *Science* **219**, 510 (1983).
- Knudson, A.G. *Am. J. hum. Genet.* **31**, 401 (1979).
- Balaban-Malenbaum, G. *et al. Cancer Genet. Cytogenet.* **3**, 243 (1981).
- Erikson, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 820 (1983).
- Shen-CKe, G.L.C., Keith, E., Piccoli, S. & Cole, M. *Cell* **31**, 443 (1982).
- Benedict, W. *et al. Science* **219**, 973 (1983).
- Rowley, J.D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5729 (1977).
- Tabin, C.J. *et al. Nature* **300**, 143 (1982).
- Reddy, E.P., Reynolds, R.K., Santos, E. & Barbacid, M. *Nature* **300**, 149 (1982).

Astronomy

The Geminga enigma unfolds

from A. W. Strong

THE γ -ray source CG195+04 ('Geminga') discovered by the SAS-2 satellite and observed repeatedly by COS-B is the second brightest source in the COS-B catalogue; unlike the other two bright sources (the Vela and Crab pulsars) it has resisted many attempts to find a convincing counterpart at other wavelengths. At a meeting held by the Royal Astronomical Society in May, G. Bignami (CNR, Milan) reported observations made in collaboration with P. Caraveo (CNR, Milan) and R. Lamb (Iowa State University) that suggest they have been able to identify X rays from the same source, and may at last be on the way to resolving the enigma.

The basic problem has always been the large radius of the γ -ray error box (radius 0.4°) which makes optical searches meaningless, there being no bright or obviously unusual object in the field. Using the imaging proportional counter (IPC) on the Einstein X-ray Observatory the authors found four sources in or near the error circle. Three are compatible with the serendipitous sample expected for the limiting flux; one can be identified with a star of typical X-ray to optical flux ratio, another with a faint (probably extragalactic) radio source, the third is unidentified. The fourth X-ray source, by far the brightest, was subsequently observed with the Einstein high-resolution imager (HRI), allowing a position to be determined to within 3 arc s. The corresponding position is blank on the Palomar sky survey; a subsequent CCD image from the European Southern Observatory revealed a blue object of about twenty-third magnitude, presumably the counterpart of the X-ray source. Meanwhile deep pulsar searches at Arecibo gave negative results, and observations with the very large array (VLA) telescope give an upper limit of 1 mJy for any source at 6-cm wavelength.

A critical clue to the physical nature of the X-ray source comes from the X-ray spectrum as determined by the IPC; it is extremely steep between 0.3 and 1 keV, which implies very little interstellar absorption (assuming a power-law or thermal spectrum at the source), and allows an upper limit on the distance of about 200 pc to be determined. The only alternative would be a line at 260 eV (with arbitrary absorption), which is also consistent with the data. No variation, periodic or otherwise, has been found for the X-ray source.

Various facts suggest the X-ray source is unique: the high X-ray to optical flux ratio ($\sim 1,000$), the point-like appearance in the HRI and the absence of a radio source in the VLA survey are properties not shared by any other Einstein source known. It is this unusual nature which makes the sug-

gestion of an association with the γ -ray source convincing, even though a strict proof is so far impossible.

The observations place severe constraints on the nature of Geminga. The X-ray to optical flux ratio rules out cataclysmic variables, white dwarfs and other hot stars. Only low-mass binary stars and radio pulsars have such high ratios, and the binary stars are ruled out because a distance of 250 kpc would be necessary to account for the faint apparent magnitude, and this is quite inconsistent with the limit from interstellar absorption as well as being far outside the Galaxy.

The remaining possibilities considered by the authors is a pulsar or just an unpulsing neutron star.

The pulsar hypothesis is attractive in view of the general similarity of the source to the Vela pulsar at X-ray and γ -ray

wavelengths, and the lack of a radio signal could in principle be assigned to different beaming geometries in γ -ray and radio wavelengths. But the lack of an extended X-ray source shows that it is not associated with a 'synchrotron nebula' of the type seen around other X-ray pulsars. The lack of a supernova remnant characteristic of fast X-ray-emitting pulsars also argues against the pulsar hypothesis.

Thus by a process of elimination they arrive at the neutron star. At a distance of 100 pc the X-ray luminosity of 10^{30} erg s $^{-1}$ implies a temperature of $2-3 \times 10^5$ K, while the best fit to the X-ray data gives 10^6 K. But neither temperature is sufficient to explain the observed optical emission, for which an additional mechanism would be needed. *A fortiori* the mechanism for the γ -ray emission, which contains more power than all other wavelengths by a factor of 1,000, and which was the start of the exercise, remains a mystery. [.]

A. W. Strong is a visiting scientist at the Istituto di Fisica Cosmica e Tecnologie Relative del CNR, Via Bassini 15, 20133 Milan.

Microbiology

Sublethal attack on bacterial membranes

from H.R. Perkins

THE membranous layers that envelope cells represent particularly sensitive areas for the destructive action of substances such as antibiotics. For agents to be therapeutically valuable they must, of course, be sufficiently selective to damage the parasite without causing undue harm to the host. An unusual example of such an apparently selective effect, confined to Gram-negative bacteria, is reported in this issue of *Nature*¹.

Bacteria are broadly classified in a binary system according to whether they stain by the procedure described in 1884 by Christian Gram² and are therefore said to be Gram-positive, or are Gram-negative. Work in the ensuing century has shown that these staining characteristics relate to fundamental differences in the organization of the bacterial envelope.

All cells necessarily have a cytoplasmic membrane with a characteristic bilayered structure containing phospholipids and numerous proteins. Outside the membrane, in which the main transport mechanisms of the cell are localized, bacteria generally have a wall layer composed of the specific amino-sugar oligopeptide co-polymer peptidoglycan. Beyond this point the Gram-negative cells differ, in that they have an additional highly specialized outer membrane^{3,4} that limits the passage of hydrophilic molecules of molecular weight greater than about 600

and has specific uptake proteins for many important substances. This outer membrane has the peculiarity that whereas the inner layer of its lipid leaflet contains phospholipids rather like those of the cytoplasmic membrane, the outer lipid layer consists of lipopolysaccharide (LPS)⁵.

It is the LPS in the Gram-negative bacterial membrane that seems to be the focus of the initial attack by the membrane-active antibiotic polymyxin B, which consists of a basic cyclic-peptide portion and a short peptide side chain substituted at its N-terminus with a nine-carbon fatty acid. Polymyxin B will interact with bacterial membranes containing LPS to yield blebs visible under the electron microscope⁶, and in a recent survey of 35 bacterial species from 27 genera, Wiegand and Quandt⁷ have shown that only bacteria that are biologically 'Gram-negative' will give blebs. (This relationship held even when the results of the Gram stain itself were difficult to interpret, so the procedure should prove very useful in the study of newly isolated bacterial species.)

Nonetheless, the action of polymyxins is not strictly specific and they are toxic to mammalian cells. This toxicity can be reduced by the removal of the fatty acid substituent, but the molecule then also loses its bactericidal activity⁸.

In this issue of *Nature*, Vaara and Vaara¹ have followed up this research by asking

whether the residual oligopeptide retains its membrane-disruptive effect. Their results are striking. The residual polymyxin B nonapeptide (PMBN) was supplied along with a series of antibiotics known to be excluded by the intact outer membrane of a smooth encapsulated clinical strain of *Escherichia coli*. In all cases the PMBN (itself non-bactericidal) decreased the minimum bactericidal concentration of the other antibiotics, the most striking examples being a 30-fold improvement in the effect of erythromycin caused by only 0.3 $\mu\text{g ml}^{-1}$ of PMBN and a 100-fold improvement in the effect of rifampicin caused by 3 $\mu\text{g ml}^{-1}$ of PMBN. Similar results were obtained with several other Gram-negative bacteria that cause severe infections, so long as those bacteria were not also polymyxin resistant.

Vaara and Vaara took their investigation further. If PMBN damaged cells and admitted antibiotics, might it not also promote the lytic action of complement? Once more the results were dramatic. Low concentrations of PMBN rendered sensitive to guinea-pig complement smooth encapsulated strains of *E. coli* and *Salmonella typhimurium* that were otherwise resistant. The explanation they offer is that the nonapeptide interacts with the outer membrane to expose hydrophobic aspects normally hidden, thereby permitting the insertion of the membrane-attacking hydrophobic complex of complement⁹.

Thus PMBN acts both to facilitate the penetration of antibiotics and to permit destruction of the bacteria by complement. Its possible value in the treatment of infections is obvious.

As so often happens, PMBN is also likely to find value as a microbiological tool. For example, ether treatment of Gram-negative cells¹⁰ has been widely used to permit the penetration of nucleotide precursors for *in vitro* synthesis of DNA¹⁰ or peptidoglycan¹¹, and no doubt PMBN will be investigated as an alternative mild treatment. Similarly, it could prove valuable in facilitating the study of periplasmic enzymes such as β -lactamases. □

H.R. Perkins is Professor of Microbiology at the University of Liverpool, PO Box 147, Liverpool L69 3BX.

Neurophysiology

Plasticity in the visual cortex

from Adam M. Sillito

EVER since Wiesel and Hubel¹ first demonstrated the striking changes in the organization of the cat visual cortex following monocular deprivation in early life, the visual system has been extensively used as a model for studying developmental plasticity in the mammalian brain.

There is common agreement throughout the literature² that during a 'critical period' extending from the time just after eye-opening to three months of age, the functional organization of the cat visual cortex is uniquely susceptible to modification by the visual environment. Monocular deprivation produced by lid suture in kittens within this period results in a drastic contraction of the deprived eye's sphere of influence in the cortex and a compensatory expansion of that of the normal eye. Periods of deprivation of a week or less are effective in eliciting the changes; and in reverse-suturing experiments, alternating the deprivation of the two eyes results in the domination of the cortex first by one eye and then by the other³. These alternations in functional domination of the cortex suggest a remarkable capacity for change, and yet outside the critical period even a year or more of monocular deprivation has little or no effect. This immediately raises the question: can any specific influence account for the period of plasticity? If so can a period of brain plasticity be reinvoked experimentally?

Our first insight into an influence that might contribute to the plasticity seen in the critical period came from the work of Kasamatsu and Pettigrew^{4,5}. They followed up the suggestion of Crow⁶ that brain stem monoaminergic systems may be involved in forebrain plasticity. Using intraventricular injection of 6-hydroxydopamine (6-OHDA), a specific neurotoxin taken up by nerve terminals containing the catecholamines noradrenaline and dopamine, they observed that animals so treated within the critical period no longer responded to the effects of monocular deprivation. This loss of plasticity was attributed to a destruction of the noradrenergic input to the visual cortex from the locus coeruleus. However, the intraventricular injections of 6-OHDA had widespread effects on brain function and so the specificity of action was open to question. This possibility was countered in subsequent experiments⁷ using a microperfusion technique to treat a localized region of the visual cortex with 6-OHDA. The result was a patch of visual cortex in the treated kittens that was resistant to the effects of monocular deprivation. Moreover, microperfusion with noradrenaline restored plasticity in 6-OHDA-treated kittens and partially restored it in older animals.

Without doubt Kasamatsu's and Pettigrew's work strongly implicates the noradrenergic input in cortical plasticity; but several recent papers urge extreme caution in the extrapolation of this to any overall control of plasticity. First, in *Nature*, Bear *et al.*⁸ recently demonstrated that visual cortical plasticity survives very early destruction of the noradrenergic input, produced by systemic injections of 6-OHDA starting in the first week of life, even though comparable reductions in cortical noradrenaline levels, produced later by microperfusion in the fourth week of life, eliminated plasticity. So it seems that the timing of the 6-OHDA-mediated destruction of the noradrenergic input could be critical: if the destruction occurs very early in life, other compensatory mechanisms may come into play.

Certainly other brain regions can have a strong influence on plasticity. Singer⁹ recently demonstrated a loss of plasticity in the critical period following destruction of the medial thalamic complex. This finding was supported by further ingenious evidence suggesting that plasticity might depend on a correlation of the visual input with a 'gating input' originating from the midbrain reticular formation or the medial thalamic region¹⁰. It is also pertinent to note that changes in visual cortical organization induced by deprivation seem to be strongly dependent on other factors, such as the proprioceptive input from the extraocular eye muscles¹¹.

Accepting that the ability of the visual cortex to respond to the visual environment in the critical period may depend on a number of influences, are there still sufficient grounds for presuming that the noradrenergic input has any special part to play in cortical plasticity? There is evidence that the noradrenergic input has an influence on synaptogenesis in the visual cortex, but this seems to occur as a transient inhibitory influence on synaptogenesis restricted to the first week of life¹². It is difficult to equate the timing of this action with the known time scale of the visual deprivation experiments, where 6-OHDA treatment in the fourth week blocks cortical plasticity. There is of course the possibility that some other more subtle influence on synaptogenesis, occurring for longer periods, has gone undetected.

An alternative suggestion is that the influence of noradrenaline on plasticity in the critical period derives from its action on neuronal response properties. Ionophoretic studies¹³ have shown that noradrenaline can have both facilitatory and inhibitory effects on the responses of visual cortical cells. Recent work on hippocampal pyramidal cells^{14,15} indicates that

1. Vaara, M. & Vaara, T. *Nature* **303**, 526 (1983).
2. Gram, C. *Fortschr. Med.* **2**, 185 (1984).
3. Di Rienzo, J.M., Nakamura, K. & Inouye, M. *Rev. Biochem.* **47**, 481 (1978).
4. Osborn, M.J. & Wu, H.C.P. *A. Rev. Microbiol.* **33A**, 369 (1980).
5. Rogers, H.J., Perkins, H.R. & Ward, J.B. *Microbial Cell Walls and Membranes* (Chapman & Hall, London, 1980).
6. Schindler, P.R.G. & Teuber, M. *Antimicrob. Ag. Chemother.* **8**, 95 (1975).
7. Wiesel, J. & Quandt, L. *J. gen. Microbiol.* **128**, 2261 (1982).
8. Chihara, S. *et al. Agr. Biol. Chem.* **37**, 2455 (1973); **38**, 521 (1974).
9. Joiner, K.A., Hammer, C.H., Brown, E.J. & Frank, M.M. *J. exp. Med.* **155**, 809 (1982).
10. Vosherg, H.P. & Hoffman-Berling, H. *J. molec. Biol.* **58**, 739 (1971).
11. Mirelman, D., Yashous-Gan, Y. & Schwarz, U. *Biochemistry* **15**, 1781 (1976).

facilitation results from a block of the calcium-dependent potassium channel that underlies the slow after-hyperpolarization that follows excitation due to the influx of sodium or calcium. The effect of this is to potentiate greatly the response of pyramidal cells to depolarization.

Now if this effect of noradrenaline were involved in plasticity it would be consistent with a hypothesis put forward by Singer¹⁰ following from Hebb's work. Singer's central thesis is that synaptic modification occurs only when dendritic depolarization of the postsynaptic target cells passes a critical level and that this level requires the conjunction of a 'gating input' with the visual input. Clearly the proposed facilitatory action of noradrenaline would greatly potentiate the postsynaptic response to a visual input. But so also, for example, would that of acetylcholine, which seems to influence both the voltage- and calcium-dependent potassium channels in pyramidal cells^{16,17}. We know that the visual cortex receives a cholinergic input, so in the light of present evidence, it seems essential to check the involvement of cholinergic processes in

plasticity. Until this has been done, both for acetylcholine and for other substances with similar actions, any unique role of noradrenaline in relation to visual cortical plasticity must stand open to question. []

Adam M. Sillito is Professor in the Department of Physiology, University College, Cardiff CF1 1XL.

1. Wiesel, T.N. & Hubel, D.H. *J. Neurophysiol.* **26**, 1003 (1963).
2. Barlow, H.B. *Nature* **258**, 199 (1975).
3. Blakemore, C. & Van Slyters, R.C. *J. Physiol., Lond.* **237**, 195 (1974).
4. Kasamatsu, T. & Pettigrew, J.D. *Science* **194**, 206 (1976).
5. Kasamatsu, T. & Pettigrew, J.D. *J. comp. Neurol.* **185**, 139 (1979).
6. Crow, T.J. *Nature* **219**, 736 (1968).
7. Kasamatsu, T. & Pettigrew, J.D. *J. comp. Neurol.* **185**, 163 (1979).
8. Bear, M.F. *et al. Nature* **302**, 245 (1983).
9. Singer, W. *Expl Brain Res.* **47**, 209 (1982).
10. Singer, W. *Expl Brain Res.* **47**, 223 (1982).
11. Freeman, R.D. & Bonds, A.B. *Science* **206**, 1093 (1979).
12. Blue, M.I. & Parnavelas, J.G. *J. comp. Neurol.* **205**, 199 (1982).
13. Kasamatsu, T. & Heggeland, P. *Expl Brain Res.* **45**, 317 (1982).
14. Madison, D.V. & Nicoll, R.A. *Nature* **299**, 636 (1982).
15. Hass, H.L. & Konnerth, A. *Nature* **302**, 432 (1983).
16. Halliwell, J.V. & Adams, P.R. *Brain Res.* **250**, 71 (1982).
17. Bernado, L.S. & Prince, D.A. *Brain Res.* **249**, 315 (1982).

Multicomponent convection

Turbulence in Earth, Sun and sea

from Herbert E. Huppert

EVERYONE is familiar with one-component convection, such as thermal convection, which occurs when pure water is heated in a kettle. But when, as in the oceans, there are two components — heat and salt — with hot salty water overlying colder fresher water, strong convective flows can be driven by the vertical salinity gradient even though dense fluid everywhere underlies less dense fluid and naive arguments would predict stability. Less than twenty years ago such multicomponent convection was considered merely an 'oceanographical curiosity'. Today it is realized that in many disciplines, complex convective patterns can arise in fluids with two or more components of different molecular diffusivities. In a chemical experiment the two components may be different inorganic solutes or high-molecular-weight polymers. Inside a star there may be four components — angular momentum, heat, the magnetic field and the hydrogen/helium composition — all diffusing at different rates. The broad implications of multicomponent convection were illustrated at a recent conference* where astrophysicists, chemists, fluid dynamicists, geologists, geophysicists, metallurgists and solar pond engineers met to discuss ideas and results relevant to their fields.

S. Turner (Australian National University) described experiments illustrating two main points. First, that double-diffusive

transport can be much larger than either the heat or mass transport in a single-component fluid, and second, the need to consider two-dimensional effects due to property gradients in the horizontal as well as the vertical. Application of the results to problems as diverse as the storage of liquid natural gas, the evolution of stars and the melting of icebergs was described. When effects due to the cooling and crystallizing of liquids are also incorporated, the subject is of considerable fundamental fluid mechanical interest, as well as opening up a new area of geology — the fluid dynamical evolution of magma chambers beneath volcanoes. The similarities between the new results on the melting of a vertical ice wall and the solidification on the side of a magma chamber were described, with the concluding question: could igneous geologists be expected to read papers on melting icebergs?

The underlying linear and nonlinear theories were reviewed by H. Huppert and M. Proctor (University of Cambridge). They contrasted the situation when linear motion sets in as a monotonic mode — as can occur, for example, when hot salty water overlies colder fresher water — with that for which an oscillation is initiated — as occurs when the colder fresher layer is on top. In the former case the resulting nonlinear motions are not very different from the linear ones, while in the latter case strong nonlinear effects can occur quite close to the linearly marginal state. For example, the formation of a nonlinear wave

packet in a horizontally unbounded domain, which breaks up and develops chaotic motion, was demonstrated in a computer simulation at parameter values close to the linear oscillatory bifurcation point (C. Bretherton, Massachusetts Institute of Technology). Further from this bifurcation point, numerical integrations of the partial differential equations for nonlinear double-diffusive convection in two-dimensional rolls between horizontal boundaries have solutions which reproduce the period-doubling behaviour and transition to chaos found for many ordinary differential and difference equations. An example of a chaotic solution is shown in Fig. 1.

Further applications were described in connection with solar ponds, which can be as large as 10 acres $\times$ 4 m deep and which trap heat from the Sun in a uniformly hot very salty 3-m storage region beneath a region with strong gradients in temperature and salinity. A successful solar pond is one in which there is 'double-diffusive non-convection'. This is because any convection in the gradient region tends to break it up into a series of layers, the hallmark of multicomponent convection. This greatly increases the heat loss from the storage region to the atmosphere, hence decreasing the efficiency of the pond (F. Zangrando, Solar Energy Research Institute, Golden, Colorado). These observations have led to attempts to calculate the temperature and salinity gradients necessary to inhibit convection (L. Bertram, Sandia National Laboratories, Albuquerque, and I. Walton, Imperial College, UK).

Double-diffusive convection has sometimes been regarded as a nuisance to chemists measuring diffusion coefficients. Alternatively, however, the special effects can be exploited to make valuable measurements, such as has been done for the Soret coefficient of NaCl (D. Caldwell, Oregon State University). Furthermore, eliminating double-diffusive effects from experiments might make the measurements irrelevant to the more usual cases in which convection will be initiated. The importance of the cross-diffusion terms in ternary or higher-order systems was clearly illustrated by the formation of ordered finger-like structures in a wide variety of polymer systems in which the concentration of neither polymer increased with height (B. Preston, Monash University, Australia, and J. Wells, Uppsala University, Sweden). A theory has recently been given to account for this effect (McDougall & Turner *Nature* **299**, 812; 1982), but it is not yet clear whether the theory accurately describes the experimental results.

Observations of magnetic fields at the surface of the Sun indicate that flux is confined to isolated tubes with diameters varying from tens of thousands of kilometres in sunspots to a hundred kilometres for fields between granules. Such fields are examples of structures produced by strongly nonlinear multicomponent convection and

*A conference on 'Double-diffusive convection' was held in Santa Barbara, under the auspices of the United States Engineering Foundation on 13–18 March 1983.

have been extensively modelled (N. Weiss, University of Cambridge). Deep in the Sun and in other stars, strong motions may be driven either by unstably stratified magnetic fields or by radial variations in the rotation rate. So far, theories to account for such flows have necessarily been simple, and mainly linear and axisymmetric. Lack of observations also constrains theoretical calculations of the convection in the liquid core of the Earth, which is believed to be the cause of the Earth's magnetic field and its aperiodic reversals. One wonders about the influence of intense motions isolated in either space or time. Are there any important analogues in the interiors of the Sun or stars or at the boundaries of the Earth's liquid core of either the metre-wide vents at the bottom of the ocean from which water around 350°C emanates with a velocity of metres per second or the strong temperature gradients observed in the oceans, such as across fronts?

Oceanographic observations which provided the motivation for much of the early development of double-diffusive convection also attest to its influence. It still remains to determine quantitatively just how important double-diffusive effects are to oceanographic mixing and to develop a clear signature to distinguish double-diffusive effects from those of internal-wave mixing, turbulence and the like.

Parts of igneous geology have been given a new lease of life by the recent rash of fluid dynamical papers devoted to exploring multicomponent convection in evolving magma chambers. An obvious place to apply and test the new concepts is the Skaergaard intrusion, a very large layered complex in east Greenland. The idea that the Skaergaard layering is due to the settling of crystals after they form has a long history but is incompatible with the measured densities of crystals and the melts from which they originate. The layers look like double-diffusive layers, but the correct quantitative signature of such layers in a crystallizing system is still being actively sought.

The often unwanted dendritic growth of a crystallizing interface is well known to metallurgists. Some dendrites are due to an instability of the originally planar interface as it grows upwards into an increasing temperature field and a decreasing concentration field set up as relatively heavy components in the melt are rejected by the solid phase. A series of experiments with solidifying succinonitrile (N. Singh, Rensselaer Polytechnic Institute, New York) showed that the formation of dendrites can retard the volume growth by as much as a factor of seven when compared with the theory for a planar interface. A theory for the rate of advance of a parabolic dendrite tip is in better agreement with the experiments, until the main dendrite faces become unstable to secondary branching and produce smaller dendrites emanating at right angles to the original one. It would therefore ap-

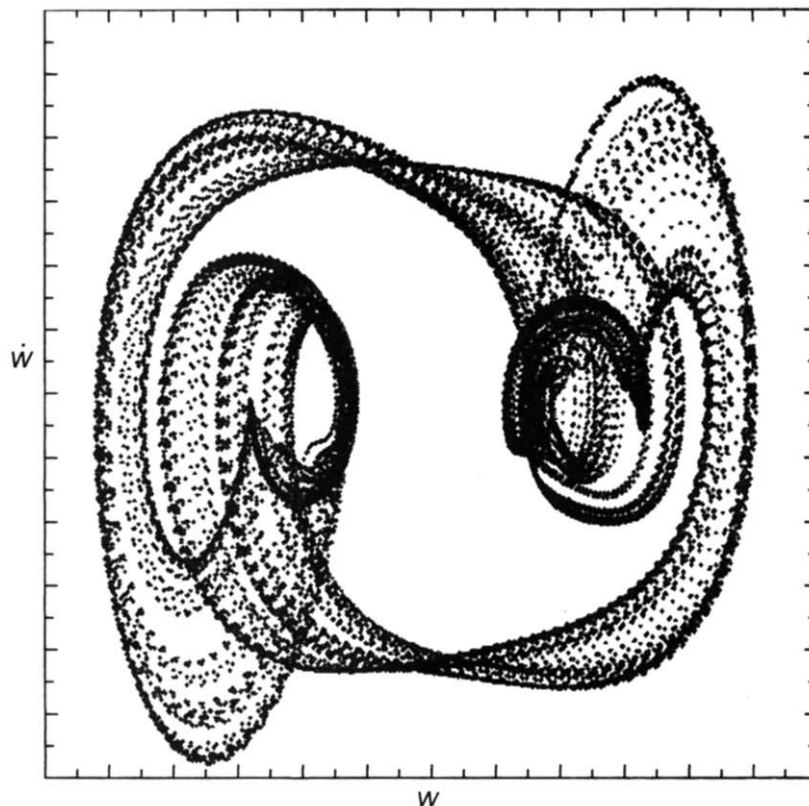


Fig. 1 Chaos in double-diffusive convection between horizontal boundaries. The phase plane of w , the vertical velocity at the edge of a convection cell mid-way between the boundaries, against $\dot{w}$, its time derivative. The solution at each time step in the numerical integration of the partial differential equations is represented by a dot in the phase plane.

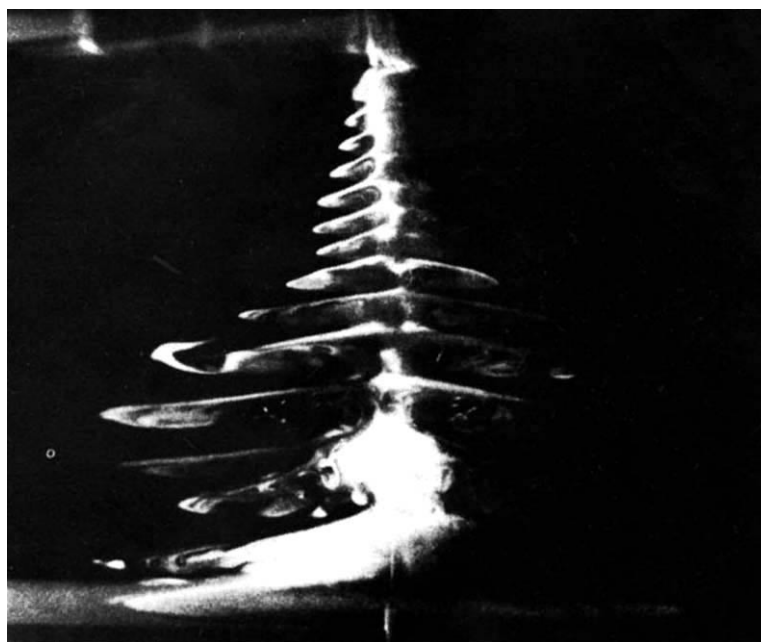


Fig. 2 Flow pattern produced by a point source of heat at the base of a stable salinity gradient. Dye at the level of the heat source has been lifted up and spread out in the 'Christmas-tree' pattern by the induced convection.

pear that further theory is still required.

In the session on laboratory experiments, A. Tsinober (Tel-Aviv University, Israel) described the effects of heating a compositional gradient from a point source. A Christmas-tree pattern of evolving layers develops (Fig. 2), with large vertical and horizontal variability. It is easy to

imagine this fundamental process, which was identified in the laboratory, occurring in many different natural circumstances. □

Herbert E. Huppert is at the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW.

Infrared astronomy

First list of IRAS sources

from the IRAS working group

THE Infrared Astronomical Satellite (IRAS) was launched on 26 January 1983 and is operating beautifully. It is expected to have a lifetime of as long as eleven months. It spends about 60 per cent of the time performing a survey of the entire infrared (IR) sky.

The IRAS team plans to publish a complete catalogue of the sky survey within one year of the end of the mission. Meanwhile, it will publish circulars about every two to three weeks and the first is presented here.

Each circular will contain from 10 to 50 sources selected for a variety of reasons.

The most important are: (1) the objects are strong emitters of far-IR radiation; (2) non-IRAS observations are critical for proper interpretation; (3) the objects are typical or interesting representatives of certain classes of astronomical objects.

The instruments in the focal plane have been described elsewhere (*J. Br. interplanet. Soc.* 36, no. 1; *Nature* 303, 287; 1983). Most observations made by the survey instrument are in four broad-wavelength bands whose parameters are given in the table in the next column.

Source name IRAS	Coordinates RA/dec	Flux densities (Jy) at				Remarks
		12 μ m	25 μ m	60 μ m	100 μ m	
0344+327P01	03h 44m 32s +32° 42.5'	1.6	5.2	18	22	Dark cloud; B5 close to CO peak
0401+261P01	04h 01m 40s +26° 10.8'	3.3	16	54	75	Dark cloud; between Lynds 1491 & B207
0406+085P01	04h 06m 30s +08° 31.1'	<0.3	0.4	3.7	9.2	Gal; NGC 1517/UGC 2970/ZWG 418.013
0409+054P01	04h 09m 42s +05° 25.2'	0.56	0.80	9.4	20	Gal; UGC 2983/MCG +01-11-013/ZWG 418.014
0415+014P01	04h 15m 05s +01° 26.1'	<0.2	<0.4	3.1	6.7	Gal? Between MCG +00-11-046 & MCG +00-11-047 (II ZW 007/ZWG 392.017)
0419+039P01	04h 19m 18s +03° 55.8'	<0.3	0.3	2.1	5.5	Gal; IC 2057/ZWG 419.002
0426+647P01	04h 26m 02s +64° 44.4'	1.0	8.0	50	57	Gal; NGC 1569
0432-143P01	04h 32m 33s -14° 19.2'	1.1	2.8	7.9	10	Dark cloud; Lynds 1642/Gal; MCG-02-12-042?
0459-341P01	04h 59m 50s -34° 06.1'	<0.2	0.4	2.9	5.3	Compact group of 3 gals; Klem 09/MCG-06-12-03
0505-375P01	05h 05m 59s -37° 34.5'	3.7	16	110	170	Gal; NGC 1808
0520-115P01	05h 20m 13s -11° 32.7'	<0.3	0.2	4.0	12	Interacting gals; NGC 1888/1889
1409-651P01	14h 09m 19s -65° 06.7'	19	65	280	340	Gal; A1409-65/Circinus gal
1636-487P01	16h 36m 16s -48° 65.7'	180	3.5K	>15K	>23K	Open cluster/HII region/Wolf Rayet (point-like) NGC 6193/RCW 108/WRAY 19.47
1648-591P01	16h 48m 26s -59° 08.0'	1.5	5.5	43	84	Gal; NGC 6221
1710-370P01	17h 10m 21s -37° 02.7'	32	350	890	580	Planetary nebula; NGC 6302
1755-213P01	17h 55m 05s -21° 20.8'	5.0	24	33	9.3	OH/IR star; OH 8.5 + 1.4
1827-145P01	18h 27m 40s -14° 31.2'	22	140	130	37	OH/IR star; OH 17.7-2.0
A410P01	16h 37m 44s -13° 38.1'	12	27	12	6.0	Asteroid; 410 Chloris UT 83 Feb 10.3

The source name consists of four parts: (1) 'IRAS' indicates the origin; (2) right ascension (RA) in hours and minutes; (3) declination (dec) in decimal degrees, multiplied by 10 and then truncated (e.g. +32° 42' 21" => +327); (4) an appendix starting with P and followed by the number of the circular; this appendix stresses that the data are preliminary. Position is given at Equinox 1950.0. The measurements have been made between epochs 1983.1 and 1983.3.

Wavelengths(μ m)	Detector size on sky (arc min) ²	Preliminary NEFD* (Jy)
Centre Bandwidth		
12 6	0.8 × 4.5	0.07
25 11	0.8 × 4.6	0.065
60 40	1.5 × 4.7	0.085
100 37	3.0 × 5.0	0.30

*1 sigma for a single pass.

A second instrument is a low-resolution spectrometer which will operate between 6 and 24 μ m with a resolving power between 10 at the shortest wavelengths and 35 at the longest. The spectrometer will normally measure objects brighter than 50 times the noise in the survey instrument.

The third instrument, the chopped photometric channel (CPC), produces small maps of selected objects in two bands at about 50 and 100 μ m. The instrument has a circular diaphragm of 1.2 arc min diameter and the maps are normally 3 × 3 or 9 × 9 arc min. When the CPC operates, data from the survey instrument are not recorded.

The normal survey measurements are repeated on time scales of seconds, hours and days. CPC observations are repeated at least once and more than several hours apart; by such repetitions asteroids and other rapidly moving objects can be distinguished.

A preliminary absolute photometric calibration, probably correct to within 20 per cent, has been made based on α Tauri and the asteroid Fortuna. Flux densities are given in Janskys for the central wavelengths listed above and for an energy distribution which is flat in the flux per logarithmic frequency interval. For other energy distributions, the flux density at the central wavelength will be related to that listed by: $f(\text{true}) = f(\text{flat})/K$ (true spectrum). Here K (true spectrum) is given by:

Source spectrum	12 μ m	25 μ m	60 μ m	100 μ m
4800 K black body	1.40	1.32	1.32	1.12
Flat in wavelength	0.95	0.95	0.99	0.99
Flat in flux/octave	1.000	1.000	1.000	1.000
Flat in frequency	1.09	1.07	1.05	1.02
100 K black body	1.15	0.88	1.03	1.06

The flux densities are derived assuming the sources to be point-like. Hence, if a source is extended, its integrated flux density will be greater than that quoted. The precision of the flux densities is shown by the number of significant figures quoted.

The positional uncertainties are less than 1 arc min in either coordinate and can be as small as 10 arc s. Refined estimates of the uncertainties will be given in future reports.

The circulars are the responsibility of the joint IRAS science working group of which H. Habing and G. Neugebauer are the co-chairmen. S. Pottasch (Kaptein Astronomical Institute, PO Box 800, 7900 AV Groningen, The Netherlands) will be the editor-in-charge and N. Boggess and P. Hacking (IRAS Project, Jet Propulsion Laboratory, Pasadena, California 91109) will coordinate the production of the circulars on the US side.

Laser processing of silicon

Ian W. Boyd & John I. B. Wilson

Department of Physics, Heriot-Watt University, Riccarton, Currie, Edinburgh EH14 4AS, UK

The application of laser radiation to the processing of silicon and silicon-based solid state devices has developed in the past two decades. A few of these techniques are becoming established practice in the semiconductor industry, whilst others remain the subject of research and heated debate. Techniques include the solid phase and the liquid phase regrowth of amorphous silicon layers, together with the more novel application of laser photochemistry.

BEAMS of ions, electrons and radiation are gradually replacing wet chemical processing and furnace treatments in the semiconductor manufacturing industry. At the turn of the 1970s the thermal diffusion of impurities into locally defined regions of an integrated circuit by a high-temperature furnace was gradually replaced by directed beams of energetic ions to implant the dopant to a specific depth in the silicon. Today, electron beam patterning is used for the generation of masks, which are then projected onto photosensitized layers on the silicon surface. The use of electron beams, ion beams and even X rays is being investigated for direct pattern generation with very high resolution. The most widespread laser application in this area is probably the laser scribing of whole silicon wafers into individual chips in a clean, fast, and economical process. Laser drilling can be useful when laser scribing becomes impossible and chemical etching is used as the separation method. The laser is used to drill alignment holes through the wafer, to achieve back-to-front alignment and subsequent reference marks for mask definition. Laser micromachining is also widely used for trimming thin film resistors in the manufacture of both hybrid and monolithic circuits. A technique which involves laser interaction with an inactive part of the circuit, but which improves the performance of devices, is laser gettering: by damaging the reverse side of a silicon wafer with a high-intensity pulsed laser, heavy metal impurities can be drawn away from the active front surface during a high-temperature anneal.

However, it is for front surface processing that laser radiation has been most widely studied in recent years. The extremely localized and rapid heating produced by the laser can be used to alter the structure and even enhance or induce surface chemical reactions. In particular, laser annealing has received much attention.

Annealing is the essential reordering process which must follow ion-implantation. The introduction of heavy, energetic atoms into the crystalline lattice disrupts its long-range order and produces a damaged, amorphous layer. This is quite unsuitable for the operation of integrated circuits, but the crystalline integrity of the layer can be restored by a high temperature ($\sim 950^\circ\text{C}$) furnace anneal. Regrowth back to the crystalline state takes place by solid phase epitaxy (SPE), during which the implanted impurity ions are incorporated into the lattice and become electrically active electron donors or acceptors. Integrated circuits are manufactured by successive masking and ion-implantation steps to give neighbouring regions in the silicon different electrical characteristics. Lasers have been used to induce solid phase recrystallization in these ion-damaged layers, or even to regrow layers by liquid phase epitaxy (LPE). This uses the high rate at which lasers can provide energy to the lattice and has led to the discovery of other novel and associated effects, such as the formation of nonequilibrium alloys and metastable solid solutions. Laser-induced surface chemistry has also been an active area of recent research. The formation of a variety of thin films, such as metal-silicides,

polycrystalline silicon and surface oxides has been reported, and both photolytic and pyrolytic processes have been employed to laser plate, evaporate and etch Si and Si-based devices. Here we review aspects of these novel techniques, including the physical mechanisms involved and the probability of any of them becoming standard manufacturing practice in the foreseeable future.

Laser annealing

Although occasional reports appeared during the late 1960s and early 1970s laser annealing did not become popular until several years after its discovery in 1974¹. Much of the pioneering work was carried out by two Soviet groups² and later by the Italians³, and was followed between 1977 and 1978 by many reports, particularly from groups in the United States. Several international conferences devoted to this aspect of laser processing have since taken place⁴⁻⁸, whilst more than 600 papers have been published on the subject.

Two modes of laser annealing can be defined: solid-phase and liquid phase regimes, and are usually associated with continuous (cw) or pulsed laser irradiation respectively.

Solid-phase epitaxial regrowth: Although cw lasers such as the argon-ion laser can be used to melt the near surface layers of silicon, they are usually employed in a lower power scanning mode and are swept across the damaged region to produce a line of recrystallized material. This is shown in Fig. 1, where

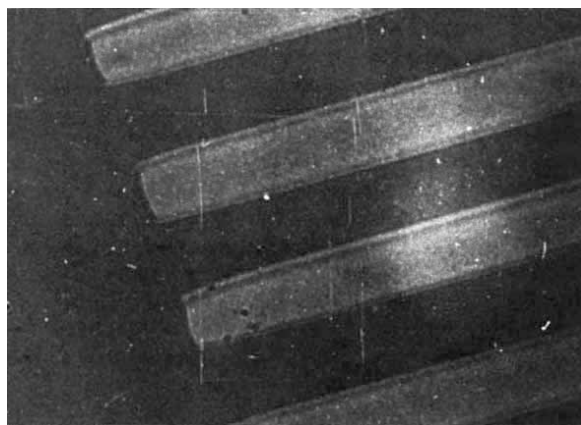


Fig. 1 Photomicrograph of cw argon laser annealed silicon. The darker region on the left is the original single crystal Si which was protected from ionic bombardment, which turned the rest of the wafer amorphous. The laser beam was then scanned across the amorphous region in a raster fashion with large spaces between neighbouring scans. On the right these neighbouring lines can be seen to have crystallized back to the same state as the original crystal Si on the left. The lighter areas are still amorphous since they were not directly exposed to the beam (from ref. 162).

successive scans have produced alternating crystalline-amorphous-crystalline regions. Typically the laser beam is scanned at speeds of $\sim 2 \text{ cm s}^{-1}$ across adjacent but overlapping regions of the material. The substrate is usually held at a temperature of $\sim 300^\circ\text{C}$ (see later) and by using beam ratios P/r (where P is power and r is $1/e$ beam radius) of $\sim 0.2 \text{ W } \mu\text{m}^{-1}$, temperatures of $\sim 1,000^\circ\text{C}$ can be achieved which are sufficient for various chemical and structural alterations. The reordering mechanism is reasonably well understood in terms of laser-solid interaction and regrowth kinetics. Although first reported in 1976⁹, it was not until later that experimental evidence favoured the popular belief that no melting took place in these conditions^{10,11}. Smoothness of surface morphology, lack of impurity diffusion and a regrowth rate of crystalline silicon of $\sim 1 \mu\text{m s}^{-1}$ at 850°C are all features of cw laser annealing.

Several models have successfully calculated the temperature profile of cw laser heated crystalline silicon¹². The rate of regrowth by SPE is limited by the time taken for the atoms to rearrange at the regrowing interface, and this requires the breaking of Si-Si covalent bonds. Therefore, the regrowth rate (R) is characterized by an Arrhenius temperature dependence, with an activation energy (E_a) close to the Si-Si bond strength, and it depends on the crystallographic orientation of the regrowing layer¹³ and the impurity concentration¹⁴. For self-implanted silicon films of (100) orientation the pre-exponent is $\sim 8 \times 10^{16} \text{ cm s}^{-1}$ and E_a is $\sim 2.4 \text{ eV}$. Laser enhanced regrowth rates of both low dose implanted silicon¹⁵ and self-implanted silicon¹⁶ have been reported, although more recent measurements have shown that there is no appreciable difference between laser and furnace regrowth rates¹⁷. (The data are summarized in Fig. 2.) The results which implied a two orders of magnitude enhancement of regrowth by laser irradiation were scaled to a melting temperature for crystalline silicon (c-Si), T_{mc} , whereas the melting temperature of amorphous silicon (a-Si), T_{ma} , has been predicted^{18,19} to be much lower: $T_{ma} \sim 0.8 T_{mc}$. Similar predictions for a-Ge have been confirmed experimentally: T_{ma} is $970\text{--}975 \text{ K}$ (refs 20, 21) compared with T_{mc} of $1,210 \text{ K}$. Unfortunately, although Baeri²² has reported a measurement of T_{ma} for a-Si of $\sim 1,170 \text{ K}$ (well below $0.8 T_{mc}$ which is $1,348 \text{ K}$), Kokorowski²³ observed an unmelted a-Si layer at temperatures in excess of $1,580 \text{ K}$. Thus, the rate at which heat is supplied to the material during these experiments is crucial.

Pulsed laser recrystallization: Pulsed laser beams may be many orders of magnitude more intense than cw beams, but only for short periods of time, with consequent differences between recrystallization in this mode and the cw annealing mechanism. For example, an amorphous layer has been estimated to regrow to a single crystal at a rate of $0.2\text{--}2.0 \text{ ms}^{-1}$ (refs 24, 25), at least four orders of magnitude greater than for SPE induced by either cw lasers or furnace methods. Furthermore, the redistribution of impurity atoms over very large distances during this annealing suggests diffusion lengths about eight orders of magnitude greater than solid state diffusion would allow during the irradiation time²⁶. This corresponds well with impurity diffusion in liquid silicon²⁷. This observation, in conjunction with the surface morphology of various samples²⁸ and the observation of a high surface reflectivity during the annealing pulse (see Fig. 3) led most early investigators to believe that melting and liquid phase epitaxy (LPE) must occur, with cooling rates as high as 10^9 K s^{-1} . Indeed, recrystallization may be prevented by using short-pulse (10 ns), short-wavelength radiation which is highly absorbed in Si; cooling rates $>10^{11} \text{ K s}^{-1}$ can be achieved, so that atoms cannot be rearranged in time to order the structure²⁹. Fast solidification can also give metastable 'frozen liquids'³⁰ in which impurity atoms cannot diffuse before the solidifying interface passes. Such nonequilibrium states can also be formed in slower regrowth conditions by zone-refining. In this case the travelling interface pushes the foreign atoms towards the surface, resulting in an extreme accumulation of dopant in the near surface region³¹. Either situation can result in 'solute trapping' where the normal solid solubility of a particular impurity in silicon is exceeded, although

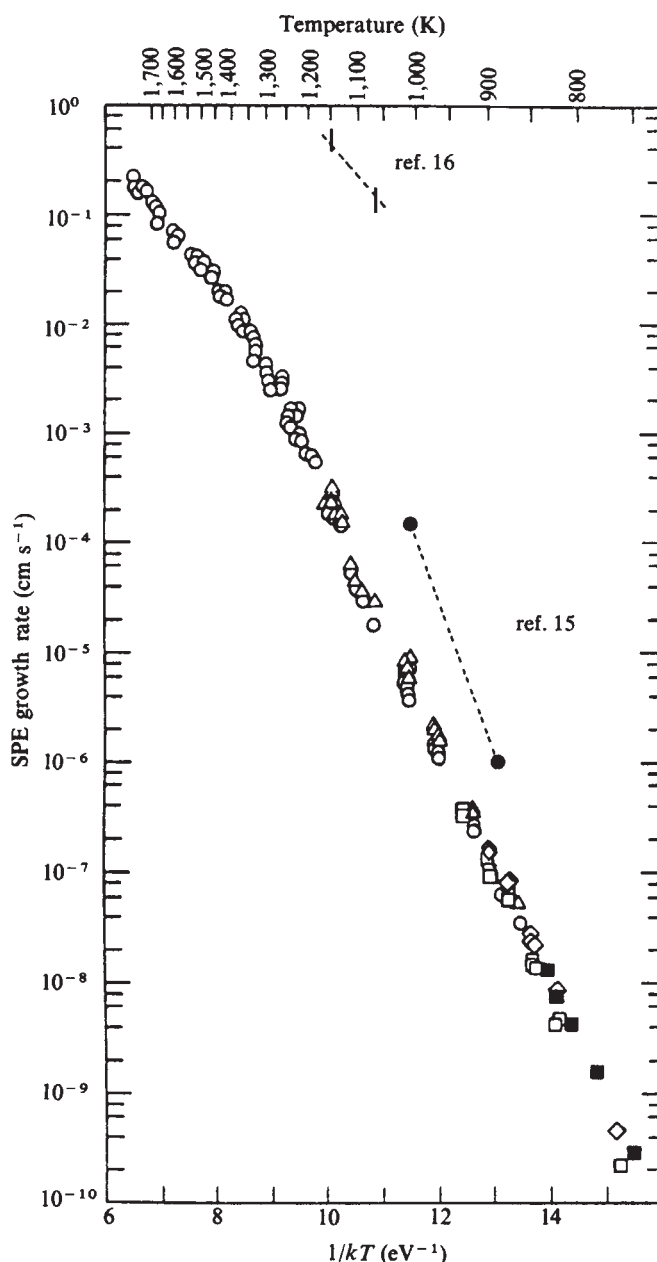


Fig. 2 Arrhenius plot comparing the various reported rates of solid-phase epitaxial regrowth of silicon produced by laser with those found from furnace annealing. Laser data are from Lietola¹⁶, Olson¹⁵ and Kokorowski, (O, Δ)¹⁷. Furnace data from Kokorowski ($\square$, $\diamond$)¹⁷ and Csepregi ($\blacksquare$)^{163,164}.

further heat treatments would relax the structure, redistributing the atoms.

The liquid phase epitaxy model can be described as follows. An amorphous layer absorbs the incident laser beam in a depth according to the absorption coefficient for that wavelength, and carriers are excited into the conduction band. The energy is then transferred from the carriers to the silicon lattice through several relaxation processes. Estimates of the carrier recombination time show that the pulse length limits the heating rate³². When the silicon melts, its optical absorption increases and the radiation is absorbed within some hundreds of angstroms at the surface. The molten layer is therefore heated further and so propagates deeper into the bulk. At the end of the laser pulse, when energy is no longer supplied to the system, the melt front halts and then retreats to the surface, as heat diffuses to the cooler surroundings. If melting occurs down to a crystal-

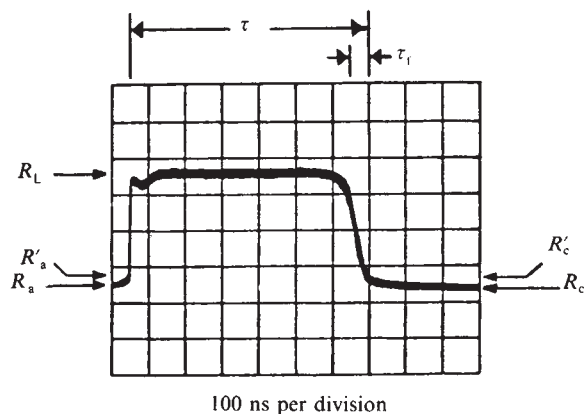


Fig. 3 Reflectivity of a He:Ne laser probe beam during and after a 30 ns Nd:YAG (532 nm) annealing pulse of 2.75 J cm^{-2} on silicon. The reflectivity starts at that for amorphous silicon (R_a), changing to that of hot, solid amorphous silicon (R'_a). After this the reflectivity jumps to that for molten silicon (R_L) for a period τ which depends on the incident energy density, then decreases to R'_c , a value ascribed to hot solid-phase crystalline silicon, during a time τ_f , finally relaxing to R_c which is that of crystal silicon at room temperature. The period τ_f is interpreted as the time taken for the final reflecting skin depth to recrystallize from the liquid phase and gives an indication of the rate of regrowth (from ref. 165).

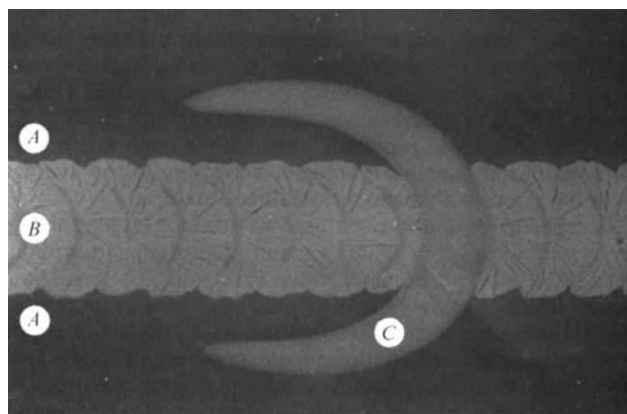


Fig. 4 Optical micrograph of amorphous silicon on glass, argon laser recrystallized by various explosive and furnace-type nucleation events. The wide band of lighter red (A) is composed of small grained ($<200 \text{ nm}$ diameter) polycrystalline Si formed by a solid-phase furnace-type crystallization mode (FCR). Within this band is a continuous line of yellow crescent-shaped features (B) which are thought to result from explosive crystallization from the liquid phase (LPXCR). The very large smooth orange crescent (C) is thought to result from solid phase XCR, and contains crystallites $\sim 500 \text{ nm}$ in diameter, as opposed to those $1 \times 10 \mu\text{m}$ observed in LPXCR. The solid-phase XCR, once nucleated at temperatures less than the melting point, will proceed such that LPXCR cannot nucleate and under appropriate power and scan conditions can also form a continuous line of XCR material. (from ref. 65).

line region then solidification occurs by LPE: the material is regrown as an extension to the crystal. If the supplied energy is insufficient for an amorphous-crystalline interface to be reached by the melt, then single-crystal seeding cannot occur, and solidification will occur at randomly distributed sites resulting in polycrystalline rather than single crystal material. In optimum conditions pulsed laser annealing can produce regrown material with the absence of such extended defects as dislocations and twinning (found in furnace annealing)²⁷.

It is widely recognized that theoretical models invoking purely thermal heating effects can accurately explain many of the phenomena associated with pulsed laser recrystallization³³⁻³⁸. Even simple models³⁹⁻⁴³ can predict the extent of dopant diffusion and melt-duration, although more elaborate models have incorporated free carrier effects^{44,45}, diffusion⁴⁶⁻⁵⁰ and even vaporization⁵¹. Nonetheless, it has still been argued that melting need not occur and that annealing may be controlled by the very high carrier densities created during the irradiation period. This alternative theory is known as the plasma annealing model^{52,53}.

In this model, energy is not immediately transferred from the excited carrier system to the lattice, but remains for relatively long times ($>10^{-7} \text{ s}$) in this state. Consequently, once a critical density of bonded electrons ($\sim 8 \times 10^{21} \text{ cm}^{-3}$) is excited, sufficient charge will have been transferred into this plasma that the covalent bonding between neighbouring atoms will be severely weakened. In such conditions a second-order phase transition could occur, in contrast with the normal first-order transition of melting. The atoms in this case would not undergo violent random vibrations as in a hot lattice, but rather would move gently into a more ordered state of lower energy. In this way, the lattice would remain relatively cool throughout the annealing process. Crucial to this argument is the mechanism by which the excited charge could remain decoupled from the lattice in such a non-equilibrium state for so long⁵⁴. The most significant and controversial data which support the plasma model are supplied by Raman measurements^{55,56} which are interpreted as revealing lattice temperatures immediately after an annealing pulse which were much lower than the melting temperature. Despite some erroneous data pertaining to observations during the high reflectivity phase shown in Fig. 3,

the results are not only still valid, but have been independently confirmed⁵⁷ using the same Raman scattering cross-sections, which, however, may be temperature dependent.

Although it seems unnecessary to invoke this model for the usual pulsed laser processing conditions, when picosecond pulses are used it is less likely that simple energy transfer from plasma to lattice is responsible for the structural changes⁵⁸. Under nanosecond or picosecond pulsed irradiation it is possible to convert crystalline silicon to amorphous silicon^{59,60} or to avoid crystallization of the complete thickness of amorphous films⁶¹ due to the rapid heating and cooling. Experiments using a modified streak camera⁶² are now underway to examine the electron diffraction pattern from samples during their irradiation by picosecond laser pulses, and these should give direct information about the structural changes during the critical period of energy transfer.

Novel processing methods

The laser offers the first means of generating extremely localized heat at a very fast rate for very short to very long times, and this novel source can be directed and moved about at will. These characteristics have led to the production of layers having hitherto unknown properties. Zone refining and solute trapping have already been discussed as the mechanisms by which supersaturated solids may be obtained, but the laser has also been used for the simple preparation and cleaning of surface layers^{63,64}. Chemical impurities previously in these regions can be removed by applying a laser pulse to the material in an ultrahigh vacuum chamber. Removal occurs by zone refinement or by diffusion into the bulk, or by vaporization or sputtering of the surface layers, and results in the formation of novel surface structures, characterized by the absence of impurities which might allow the truncated lattice to relax to a less strained configuration.

The phenomenon of explosive crystallization of amorphous layers on an insulating substrate has also been explored using cw lasers. This is a very rapid and self-sustaining reordering process in which heat liberated by crystallization is passed on to a neighbouring amorphous region to trigger further reordering and so on. Until recently, this effect had been studied by applying the initial energy required at a fixed point, but now

the laser allows a moving heat source to be used. The process is quenched when heat from the laser beam plus the heat of crystallization is insufficient. The interplay of background temperature, scan velocity and beam power can produce various patterns related to different crystallization regimes (see Fig. 4). Both solid and liquid regrowth phases have been identified⁶⁵ giving various surface morphologies according to the average size of the grains within the recrystallized material.

Laser beams have also been used to produce large grains of polycrystalline silicon from vacuum-deposited amorphous layers on crystalline or amorphous substrates, or from fine-grained layers of low pressure chemical vapour deposited (LPCVD) silicon. In selected conditions it is possible to obtain high quality single-crystal films of silicon on an insulating substrate, such as silica, which is a cheaper alternative to the developing silicon-on-sapphire (SOS) technology. The research in this area mainly concerns the scanned cw laser crystallization of CVD silicon layers or the pulsed laser crystallization of vacuum-deposited (such as electron beam evaporated) silicon.

Although cw laser annealing of electron-beam deposited silicon can produce 100- μm sized grains similar to those deposited by LPCVD, this solid phase epitaxy process is strongly affected by interface contaminants unless the film is grown in ultraclean conditions⁶⁶. Even with pulsed laser crystallization of vacuum deposited silicon, which relies on melting, the perfection of the final product is restricted by these same contaminants, although the interfacial oxide layer is less of an obstacle than it is when attempting to thermally crystallize these films (on a crystalline silicon substrate)^{67,68}.

The greatest interest is in improving the crystallinity of polycrystalline silicon by a scanned argon laser beam, for films so treated have yielded the best transistors from any non-conventional process. It is even possible to ion-implant the as-grown CVD film and combine the implant anneal with recrystallization without diffusing the dopant⁶⁹. A good technique for promoting crystal growth under the laser beam is to use photolithographically defined 'lozenge-shaped' islands of silicon having a long dimension parallel to the laser beam⁷⁰⁻⁷². This geometry can provide a single nucleation point and will prevent stresses building up in the crystallized zone. Alternatively, seeding of crystal growth can be provided by a patterned substrate, such as fine grooves in SiO_2 (refs 73, 74) or even by windows etched down to a crystalline substrate^{75, 76}. The best results also use a specially shaped beam profile to control the heat flow at the vital trailing edge of the swept irradiated area⁷⁷. If the laser energy is sufficient to melt the silicon then undesirable surface peaks and troughs may be produced⁷⁸, but absorption of weaker laser radiation may be enhanced through the addition of antireflection coatings, which may themselves be patterned to control crystal nucleation⁷⁹. These increasingly refined techniques are able to produce single crystal islands of silicon $\sim 20 \times 400 \mu\text{m}$ in area, which are large enough to be processed into solid state devices, each island electrically isolated from the others.

Laser microchemistry

Apart from laser crystallization, the term 'laser annealing' has come to include other processing techniques, such as alloying and thin film deposition. These developments can only be treated briefly here.

Laser assisted formation of metal silicides: As integrated circuits become smaller, there is a demand for very fine, highly conductive inter-connecting paths within each device. One approach involves the use of metal silicides, which inherently display much higher conductivity than the commonly used doped polysilicon. Laser beam processing opens up the possibility of producing layers not previously available with conventional heat treatments. For instance, by laser-melting a deposited metal layer such as platinum or palladium on silicon, interdiffusion of metal and silicon can take place⁸⁰. Single phase silicide layers can also be formed by solid phase reactions^{81,82}. Recent papers^{83,84} showing that good rectifying contacts for Schottky

barrier diodes and ohmic contacts of Pt-Si can be applied successfully to device manufacture are an extension of early laser alloying experiments using Al and n-Si to produce p-n junctions⁸⁵, and ohmic contacts to III-V semiconductors⁸⁶. It is significant that improvements over conventional methods of more than one order of magnitude in the contact resistance of gold-germanium (Au-Ge) contacts to gallium arsenide (GaAs) have been achieved by laser processing^{87,88}.

Laser photochemistry for Si processing: CVD of thin films can be activated by lasers so that patterns can be fabricated directly without the need for photolithography. The laser beam can either provide local substrate heating for the pyrolytic decomposition of a gaseous compound, (for example, to form silicon⁸⁹⁻⁹²), or it can directly dissociate a gaseous compound for subsequent deposition on a substrate. The latter, photolytic, process usually requires a UV laser (for example, to form Si from SiH_4)⁹³ but some gases do absorb IR radiation (for example, CO_2 laser dissociation of SiH_4 to give Si)⁹⁶⁻⁹⁸. These deposition steps can be combined with simultaneous diffusion of dopants⁹⁶⁻⁹⁸, even to produce ohmic contacts, an especially valuable alternative to the usual thermal alloying of metal films in the case of compound semiconductors which readily decompose when heated⁹⁹. It has also been possible to make p-n junction solar cells on silicon^{100,101} or GaAs¹⁰² by simultaneous gas decomposition and diffusion of the dopant so-formed. Since metal films can also be photochemically deposited from gaseous or liquid sources by a laser¹⁰³⁻¹⁰⁷ and silicon can be laser oxidized¹⁰⁸⁻¹¹⁰, there is now the possibility of growing multilayer solid state devices with semiconductor, dielectric and metal components each produced by laser beam, and avoiding conventional photolithography by steering the laser beam. Even if some chemical etching is required to define areas on the substrate, this can be assisted by laser irradiation¹¹¹⁻¹¹⁴. In all of these processes one is not limited in spatial resolution to the beam diameter, but can define submicrometre linewidths, apparently due to limited dopant diffusion¹¹⁵ or to pre-nucleation from a very thin surface absorbed layer¹¹⁶ (for doping and deposition respectively). Since semiconductor devices are being made smaller each year, and with the probability of a greater use of integrated optics for signal processing, laser chemical deposition might be the most appropriate method for directly 'writing' the required circuit elements of the newer compound semiconductors on to cheap insulating substrates.

Industrial application of laser processing

Although it is now possible by laser to recrystallize completely a disordered layer resulting from ion-implantation damage and simultaneously to incorporate virtually 100% of the implanted dose on lattice sites, the process is still not used by the integrated circuit manufacturers, although many of the original problems have been solved. For example, hotspots in pulsed laser beams which often lead to surface distortions have been virtually eliminated by Cullis¹¹⁷ through the use of a beam homogenizer to transform the incident beam into a flat-topped intensity profile. The appearance of poly-silicon 'haloes' around good quality recrystallized Si often occurs through insufficient energy density around the edge of the annealing beam to produce seeding from the single crystal sub-layer. This can be avoided by overlapping successive adjacent spots. Preheating the substrate to at least 200 °C reduces slip damage, but this also eliminates dislocation loops and stacking faults¹¹⁸. However, the appearance of high concentrations of point defects captured by the rapid quenching is common in laser annealed material. This problem can only be overcome by a further low temperature process such as heating the material in a hydrogen plasma¹¹⁹. The major remaining problem is that of non-uniform energy absorption which is particularly obvious in multimaterial systems. This makes the annealing of silicon through windows in the commonly used SiO_2 coating particularly difficult, unless further covering films of poly-silicon or aluminium are applied¹²⁰. An antireflection coating has been used to maximize the energy coupled into the required area¹²².

There are a few reports of complete devices made with laser annealing. In 1978 Young¹²³ made laser annealed diodes whose properties closely followed the ideal diode equation. In 1979 Zimmer¹²⁴ successfully used a cw laser to anneal ion-implanted channels of n-MOS transistors without degrading the operational characteristics. It has even been possible to improve the performance of some devices by incorporating a laser annealing stage^{120,121,125,126}, yet the general opinion seems to be that laser annealing does not produce devices of such consistently superior quality that costly capital equipment replacement would be justified. Additionally, the necessity of redesigning the whole set of process steps has enforced the conclusion that there is insufficient technical reason to replace furnace methods in integrated circuit manufacturing. However, if CVD poly-silicon on an insulating substrate is to be a successful semiconductor, then laser annealing is one way of obtaining good MOS transistors (see for example, refs 127–129).

Competitive annealing techniques

Inspired by the relative success of laser annealing, research groups have exploited other sources of energy for similar purposes. Foremost in these alternative annealing methods is electron-beam annealing (EBA). The feasibility of EBA was suggested by a US patent in 1974¹³⁰ even before laser annealing was discovered. The first reports of pulsed¹³¹ and cw¹³² EBA appeared in 1977 and 1979 respectively and are similar to their laser counterparts. MOS transistors have been made in electron beam recrystallized CVD poly-silicon films¹³³. The biggest disadvantages of EBA are the enormous capital outlay and the need for a vacuum chamber to allow beam propagation. In contrast, beams of electrons can be shaped to give a rectangular spot so as to allow fast scanning of whole wafers^{134,135} and coupling between the beam and substrate is independent of the degree of disorder of the target so that absorption is more homogeneous. Ion beams^{136,137} and pulsed proton beams¹³⁸ have also been used to anneal implantation damage and energetic beams of heavy ions are thought to stimulate recrystallization at low temperatures by an interaction between the amorphous-crystalline boundary and the vacancies and interstitials produced by their impact.

In contrast to this high technology, it has been demonstrated that incoherent light sources such as argon or xenon flash lamps, or even halogen lamps, can provide sufficient focused heat to anneal ion-implanted silicon at low cost, low power consumption and high speed^{139–147}. Damage recovery occurs by either SPE or LPE. Even a resistively-heated graphite strip passed over an implanted 3-inch silicon wafer can anneal it in a few seconds^{148,149}. A specially shaped beam from a halogen lamp has been used to recrystallize deposited poly-silicon on an oxidized silicon wafer, in an analogous manner to a scanned cw argon laser beam, but with considerably less expense¹⁵⁰. But the most likely commercial use of radiant beam annealing is the graphite strip-heater crystallization of deposited silicon films on insulating substrates. This can be achieved either by promoting lateral crystal growth in an amorphous or polycrystalline layer from seed regions, provided by windows etched through the insulating coating of a single crystal Si substrate or by a patterned SiO₂ substrate^{151–153}, or by dispensing with the pattern and sweeping a molten zone across the film^{154,155}.

Until localized processing methods are specifically required, beam processing will always be challenged by the traditional, more cost effective, methods of applying energy. The trend is to reduce the time that each wafer spends at elevated temperatures, but the technologies reviewed in this section have the advantage of much faster processing times than conventional furnace systems.

A promising application of pulsed laser annealing uses the large ratio of diffusion coefficients of dopants in liquid and solid silicon near the melting point. This offers the first method for completely separate optimization of the dopant profiles in the emitter and base regions of bipolar transistors, independent of

dopant, dopant concentration, or the presence of other dopants¹⁵⁶.

Pulsed lasers are also favoured for the development of simple, high-speed, automated techniques for producing low cost, high quality, p-n junction silicon solar cells. By using a combination of laser annealing of the appropriate implanted dose with laser induced diffusion and laser damage gettering, c-Si solar cells displaying efficiencies greater than 16% have been obtained, compared with ~17½% obtained on a small laboratory scale using sophisticated processing methods¹⁵⁷. Pulsed lasers have been used to drive-in vacuum-deposited dopants¹⁵⁸ or dopants from gaseous or liquid sources, in an attempt to reduce the cost of single crystal silicon p-n solar cells. cw lasers may be the better choice for annealing ion-implanted polycrystalline solar cells to avoid melting the silicon, which would redistribute the dopant along the grain boundaries of the underlying layer, and are a way of improving the crystallinity of cheaply deposited poly-silicon for solar cells¹⁵⁹. A group at Stanford University¹⁶⁰ have produced 10% efficient solar cells in low grade and polycrystalline silicon with cw argon ion laser annealing, which are better than the equivalent cells produced by furnace annealing.

At the other extreme of solid state devices where, rather than low cost, complexity is the object, it has been demonstrated that 'multilayer integration' can be achieved in laser recrystallized polysilicon¹⁶⁰. Using silicon nitride layers as masks for both laser recrystallization and, later, MOSFET definition, together with ion implantation to produce the doped layers required by such transistors, two stacked devices were made, in a form which is a building block for several logic circuits. Development of this idea could lead to more closely packed digital integrated circuits.

Conclusion

Although laser processing may not have fulfilled the promise it once was thought to have for reducing the cost of semiconductor devices, it still provides ways of making new interfaces and new materials which are otherwise unobtainable. The search for other beam processing methods has been stimulated by this work, as has the investigation of the kinetics of crystal growth and the diffusion of impurities. Both theoretical and practical aspects of materials science have been enriched by the laser processing interactions. The influence of the electron-hole plasma produced during short pulse, high energy irradiation of silicon in particular has caused theorists to speculate on new phase transitions. A new generation of semiconductor devices may evolve from the new laser photolytic and pyrolytic reactions which aid the formation of localized deposits of metallic, dielectric or semiconducting compounds without using conventional photolithographic processing.

Many people have supplied material for this review; in particular we thank Drs A. Cullis, G. Auvert, J. Van Vechten. Some of this survey was included in a thesis by I.W.B., supported by Hughes Microelectronics Ltd, and SERC.

1. Khaibullin, I. B., Shtyrkov, E. I., Zaripov, M. M., Galyautdinov, M. F. & Bajazitov, R. M. *Paper No. 2061-74*, (Viniti, Moscow, 1974 (in Russian)).
2. Khaibullin, I. B., Shtyrkov, E. I., Zaripov, M. M., Bayazitov, R. M. & Galjanudinov, M. F. *Radiat. Eff.* **36**, 225 (1978).
3. Foti, G., Rimini, E., Vitali, G. & Bertolotti, M. *Appl. Phys.* **14**, 189 (1977).
4. Ferris, S. D., Leamy, H. J. & Poate, J. M. *Proc. Mater. Res. Soc. Symp. on Laser-Solid Interactions and Laser Processing* (AIP, New York, 1979).
5. Anderson, C. L., Rozgonyi, G. A. & Celler, G. K. *Laser and Electron Beam Processing of Electronic Materials* (Electrochemical Society, 1980).
6. White, C. W. & Peercy, P. S. *Laser and Electron Beam Processing of Materials* (Academic, New York, 1980).
7. Gibbons, J. F., Hess, L. D. & Sigmon, T. W. *Laser and Electron Beam Solid Interactions and Material Processing* (North-Holland, Amsterdam, 1981).
8. Appleton, B. R. & Celler, G. K. *Laser and Electron Beam Interactions with Solids* (Elsevier, Amsterdam, 1982).
9. Kachurin, K. A., Nidaev, E. V., Khodyachikh, A. V. & Kovaleva, L. A. *Soviet Phys. Semicond.* **10**, 1128 (1976).
10. Gat, A. *et al. Appl. Phys. Lett.* **33**, 389 (1978).
11. Auston, D. H., Golovchenko, J. A., Smith, P. R., Surko, S. M. & Venkatesan T. N. C. *Appl. Phys. Lett.* **33**, 539 (1978).
12. Nissim, Y. I., Lietoila, A., Gold, R. B. & Gibbons, J. F. *J. appl. Phys.* **51**, 274 (1980).
13. Sigmon, T. W. *IEEE Trans. nucl. Sci.* **28**, 1797 (1981).
14. Lau, S. S. in *Proc. Mater. Res. Soc. Symp. on Laser-Solid Interactions and Laser Processing*, **84** (AIP, New York, 1979).

15. Olson, G. L., Kokorowski, S. A., McFarlane, R. A. & Hess, L. D. *Appl. Phys. Lett.* **37**, 1019 (1980).
16. Lietoila, A., Gold, R. B. & Gibbons, J. F. *Appl. Phys. Lett.* **39**, 810 (1981).
17. Kokorowski, S. A., Olson, G. L. & Hess, L. D. *J. Appl. Phys.* **53**, 921 (1981).
18. Bagley, B. G. & Chen, H. S. in *Proc. Mater. Res. Soc. Symp. on Laser-Solid Interactions and Laser Processing*, 97 (AIP, New York, 1979).
19. Spaepen, F. & Turnbull, D. in *Proc. Mater. Res. Soc. Symp. on Laser-Solid Interactions and Laser Processing*, 73 (AIP, New York, 1979).
20. Zeiger, H. J., Fan, J. C. C., Palm, B. J., Gale, R. P. & Chapman, R. L. *Laser and Electron Beam Processing of Materials*, 234 (Academic, New York, 1980).
21. Chapman, R. L., Fan, J. C. C., Zeiger, H. J. & Gale, R. P. *Laser and Electron Beam Solid Interactions and Materials Processing*, 81 (North-Holland, Amsterdam, 1981).
22. Baeri, P., Foti, G., Poate, J. M. & Cullis, A. G. *Phys. Rev. Lett.* **45**, 1036 (1980).
23. Kokorowski, S. A., Olson, G. L., Roth, J. A. & Hess, L. D. *Phys. Rev. Lett.* **48**, 498 (1982).
24. Auston, D. H., Golovchenko, J. A., Simons, A. L., Surko, C. M. & Venkatesan, T. N. C. *Appl. Phys. Lett.* **34**, 777 (1979).
25. Baeri, P., Compisano, S. U., Foti, G. & Rimini, E. *J. appl. Phys.* **50**, 788 (1979).
26. Wang, J. C., Wood, R. F. & Pronko, P. P. *Appl. Phys. Lett.* **33**, 455 (1978).
27. White, C. *Radiat. Eff.* **47**, 37 (1981).
28. Leamy, H. J., Rozgonyi, G. A., Sheng, T. T. & Celler, G. K. *Appl. Phys. Lett.* **32**, 535 (1978).
29. Tsu, R., Hodgson, R. T., Tan, T. Y. & Baglin, J. E. *Phys. Rev. Lett.* **42**, 1356 (1979).
30. Leamy, H. J., Bean, J. C., Poate, J. M. & Celler, G. K. *J. Cryst. Growth* **48**, 379 (1980).
31. White, C. W., Narayan, J., Appleton, B. R. & Wilson, S. R. *J. appl. Phys.* **50**, 2967 (1978).
32. Yoffa, E. J. *Phys. Rev. B21*, 2415 (1981).
33. Meyer, J. R., Barton, F. J. & Krueger, M. R. *Phys. Rev. B21*, 1559 (1980).
34. Bertolotti, M. & Sibilia, C. *IEEE J. Quant. Elect.* **QE-17**, 1980 (1981).
35. Lietoila, A. & Gibbons, J. F. *Appl. Phys. Lett.* **40**, 624 (1982).
36. Sai-Halasz, G. A. & Hodgson, R. T. *Phys. Lett.* **77A**, 375 (1980).
37. Narayan, J. & White, C. W. *Phil. Mag.* **A43**, 1515 (1981).
38. von der Linde, D. & Fabricius, N. *Appl. Phys. Lett.* **41**, 991 (1982).
39. Baeri, P., Compisano, S. U., Foti, G. & Rimini, E. *J. appl. Phys.* **50**, 788 (1979).
40. Bell, A. E. *RCA Rev.* **40** (1979).
41. Schultz, J. C. & Collins, R. J. *Appl. Phys. Lett.* **34**, 84 (1979).
42. Bell, R. O., Toulemonde, M. & Siffert, P. *Appl. Phys.* **19**, 313 (1979).
43. Meyer, J. R., Krueger, M. R. & Bartoli, F. J. *J. appl. Phys.* **51**, 5513 (1980).
44. Lietoila, A. & Gibbons, J. F. *Appl. Phys. Lett.* **34**, 332 (1979).
45. Nilsson, N. G. & Svantesson, K. G. *J. Phys. D13*, 39 (1980).
46. Wood, R. F. *Appl. Phys. Lett.* **38**, 357 (1981).
47. Jastrzebski, L., Bell, A. E., Wu, P. & Zanzucchi, P. J. *J. appl. Phys.* **52**, 4276 (1981).
48. Kim, D. M. & Kwong, D. L. *IEEE J. Quant. Electron.* **18**, 224 (1982).
49. Meyer, J. R., Barton, F. J. & Krueger, M. R. *Phys. Rev. B21*, 1559 (1980).
50. Wood, R. F. & Giles, G. E. *Phys. Rev. B23*, 2923 (1981).
51. Jain, A. K., Kulkarni, V. N. & Sood, D. K. *Appl. Phys.* **25**, 127 (1981).
52. Van Vechten, J. A., Tsu, R., Saris, F. W. & Hoonhout, D. *Phys. Lett.* **74**, 417 (1979).
53. Van Vechten, J. A., Tsu, R. & Saris, F. W. *Phys. Lett.* **74**, 422 (1979).
54. Phillips, J. C. *J. appl. Phys.* **52**, 7397 (1981).
55. Lo, H. W. & Compain, A. *J. appl. Phys.* **51**, 1565 (1980).
56. Lo, H. W. & Compain, A. *Phys. Rev. Lett.* **44**, 1604 (1980).
57. von der Linde, D. & Wartmann, G. *Appl. Phys. Lett.* **41**, 700 (1982).
58. Merkle, K. L., Baumgart, H., Uebbing, R. H. & Philipp, F. *Appl. Phys. Lett.* **40**, 729 (1982).
59. Liu, P. L., Yen, R., Bloembergen, N. & Hodgson, R. T. *Appl. Phys. Lett.* **34**, 864 (1979).
60. Cullis, A. G., Webber, H. C., Chew, N. G., Poate, J. M. & Baeri, P. *Phys. Rev. Lett.* **49**, 219 (1982).
61. Liang, P. H., Fang, C. J., Jiang, D. S., Wagner, P. & Ley, L. *Appl. Phys. A26*, 39 (1981).
62. Mourou, G. & Williamson, S. *Appl. Phys. Lett.* **41**, 44 (1982).
63. Zehner, D. M., White, C. W., Ownby, G. W. *Appl. Phys. Lett.* **36**, 56 (1980).
64. Cowan, P. L., Golovchenko, J. A. *J. Vac. Sci. Technol.* **18**, 1197 (1980).
65. Auvert, G., Bensahel, D., Peris, A., Nguyen, V. T. & Rozgonyi, G. A. *Appl. Phys. Lett.* **39**, 724 (1981).
66. Roth, J. A., Olsen, G. L., Kokorowski, S. A. & Hess, L. D. *Laser and Electron Beam Solid Interactions and Materials Processing*, 413 (North-Holland, Amsterdam, 1981).
67. Lau, S. S. et al. *Appl. Phys. Lett.* **33**, 130 (1978).
68. Celler, G. K. et al. *Laser and Electron Beam Solid Interactions and Materials Processing*, 435 (North-Holland, Amsterdam, 1981).
69. Gat, A. et al. *Appl. Phys. Lett.* **33**, 775 (1978).
70. Gibbons, J. F., Lee, K. F., Magee, T. J., Peng, J. & Ormond, R. *Appl. Phys. Lett.* **34**, 831 (1979).
71. Johnson, N. M., Biegelsen, D. K. & Moyer, M. D. *Laser and Electron Beam Solid Interactions and Materials Processing*, 463 (North-Holland, Amsterdam, 1981).
72. Biegelsen, D. K., Johnson, N. M., Bartelink, D. J. & Moyer, M. D. *Laser and Electron Beam Solid Interactions and Materials Processing*, 487 (North-Holland, Amsterdam, 1981).
73. Geis, M. W., Flanders, D. C. & Smith, H. I. *Appl. Phys. Lett.* **35**, 71 (1979).
74. Tamura, M., Tamura, H. & Tokuyama, T. *Jap. J. appl. Phys.* **19**, L-23 (1980).
75. Magee, T. J., Palkuti, L. J., Ormond, R., Leung, C. & Graham, S. *Appl. Phys. Lett.* **38**, 248 (1981).
76. Celler, G. K., Trimble, L. E., Ng, K. K., Leamy, H. J. & Baumgart, H. *Appl. Phys. Lett.* **40**, 1043 (1982).
77. Kawamura, S., Sakurai, J., Nakano, M. & Takagi, M. *Appl. Phys. Lett.* **40**, 394 (1982).
78. Iso, T., Sugahara, K., Nishimura, T., Akasaka, T. & Nakata, H. *Appl. Phys. Lett.* **40**, 788 (1982).
79. Colinge, J. P., Demoulin, E., Bensahel, D. & Auvert, G. *Appl. Phys. Lett.* **41**, 346 (1982).
80. Poate, J. M., Leamy, H. J., Sheng, T. T. & Celler, G. K. *Appl. Phys. Lett.* **33**, 918 (1978).
81. Gat, A. et al. *Appl. Phys. Lett.* **32**, 276 (1978).
82. Liau, Z. L. *Appl. Phys. Lett.* **34**, 221 (1979).
83. Yang, E. S. & Yu, C. M. *Appl. Phys. Lett.* **37**, 462 (1980).
84. Doherty, C. J., Seidel, T. E., Leamy, H. J. & Celler, G. K. *J. appl. Phys.* **51**, 2718 (1980).
85. Harper, F. E. & Cohen, M. I. *S.S. Elect.* **13**, 1103 (1970).
86. Pounds, R. S., Saifa, M. A. & Hahn, W. C. *S.S. Elect.* **17**, 245 (1974).
87. Barnes, F. A. et al. *Appl. Phys. Lett.* **33**, 965 (1978).
88. Oraby, A. H., Murakami, K., Yuba, Y., Gamo, K. & Namba, S. *Appl. Phys. Lett.* **38**, 562 (1981).
89. Christensen, C. P. & Lakin, K. M. *Appl. Phys. Lett.* **32**, 254 (1978).
90. Baranaskas, V., Mammana, C. I. Z., Klinger, R. E. & Green, J. E. *Appl. Phys. Lett.* **36**, 930 (1980).
91. Ehrlich, D. J., Osgood, R. M. & Deutsch, T. F. *Appl. Phys. Lett.* **39**, 957 (1981).
92. Bäuele, D., Irsigler, P., Leyendecker, G., Noll, H. & Wagner, D. *Appl. Phys. Lett.* **40**, 819 (1982).
93. Andreatta, R. W. et al. *Appl. Phys. Lett.* **40**, 183 (1982).
94. Hanabusa, M., Namiki, A. & Yoshihara, K. *Appl. Phys. Lett.* **35**, 626 (1979).
95. Bilenchi, R., Gianinoni, I. & Musci, M. *J. appl. Phys.* **53**, 6479 (1982).
96. Ehrlich, D. J., Osgood, R. M. & Deutsch, T. F. *Appl. Phys. Lett.* **36**, 916 (1980).
97. Deutsch, T. F., Ehrlich, D. J., Rathman, D. D., Silversmith, D. J. & Osgood, R. M. *Appl. Phys. Lett.* **39**, 825 (1981).
98. Ibbes, K. G. & Lloyd M. L. *Opt. Laser Technol.* **35** (1983).
99. Deutsch, T. F., Ehrlich, D. J., Osgood, R. M. & Liau, Z. L. *Appl. Phys. Lett.* **36**, 847 (1980).
100. Deutsch, T. F. et al. *Appl. Phys. Lett.* **38**, 144 (1981).
101. Turner, G. B., Tarrant, D., Pollock, G., Pressley, R. & Press, R. *Appl. Phys. Lett.* **39**, 967 (1981).
102. Deutsch, T. F. et al. *Appl. Phys. Lett.* **40**, 722 (1982).
103. Deutsch, T. F., Ehrlich, D. H., & Osgood, R. M. *Appl. Phys. Lett.* **35**, 175 (1979).
104. Ehrlich, D. J., Osgood, R. M., Silversmith, D. J. & Deutsch, T. F. *IEEE Electron Dev. Lett.* **EDL-1**, 101 (1980).
105. Solanki, R., Boyer, P. K., Mahan, J. E. & Collins, G. J. *Appl. Phys. Lett.* **38**, 572 (1981).
106. Karlicek, R. F., Donnelly, V. M. & Collins, G. J. *J. appl. Phys.* **53**, 1084 (1982).
107. Tsao, J. Y., Ehrlich, D. J., Silversmith, D. J. & Mountain, R. W. *IEEE Electron Dev. Lett.* **EDL-3**, 164 (1982).
108. Boyd, I. W., Wilson, J. I. B. & West, J. L. *Thin Solid Films* **83**, L173 (1981).
109. Boyd, I. W. & Wilson, J. I. B. *Appl. Phys. Lett.* **41**, 162 (1982).
110. Boyer, P. K., Roche, G. A., Ritchie, W. H. & Collins, G. J. *Appl. Phys. Lett.* **40**, 716 (1982).
111. Ehrlich, D. J., Osgood, R. M. & Deutsch, T. F. *Appl. Phys. Lett.* **36**, 698 (1980).
112. Haynes, R. W., Metze, G. M., Kreismanis, V. G. & Eastman, L. F. *Appl. Phys. Lett.* **37**, 344 (1980).
113. Ehrlich, D. J., Osgood, R. M. & Deutsch, T. F. *Appl. Phys. Lett.* **38**, 1018 (1981).
114. Osgood, R. M., Sanchez-Rubio, A., Ehrlich, D. J. & Daneu, V. *Appl. Phys. Lett.* **40**, 391 (1982).
115. Ehrlich, D. J. & Tsao, J. Y. *Appl. Phys. Lett.* **41**, 297 (1982).
116. Ehrlich, D. J., Osgood, R. M. & Deutsch, T. F. *Appl. Phys. Lett.* **38**, 946 (1981).
117. Cullis, A. G., Webber, H. C. & Bailey P. J. *Phys. (E): Sci. Instrum.* **12**, 688 (1979).
118. Celler, G. K., Poate, J. M., Rozgonyi, G. A. & Sheng, T. T. *J. appl. Phys.* **50**, 7264 (1979).
119. Benton, J. L. *Appl. Phys. Lett.* **36**, 338 (1980).
120. Koyanagi, M., Tamura, H., Miyao, M., Hashimoto, N. & Tokuyama, T. *Appl. Phys. Lett.* **35**, 621 (1979).
121. Hill, C. *Laser and Electron Beam Processing of Electronic Materials*, 26 (Electrochemical Society, 1980).
122. Teng, T. C., Shiau, Y., Chen, Y. S. & Skinner, C. *Laser and Electron Beam Solid Interactions and Materials Processing*, 391 (North-Holland, Amsterdam, 1981).
123. Young, R. T., Narayan, J. & Wood, R. F. *Appl. Phys. Lett.* **35**, 447 (1979).
124. Zimmer, G. *Electron. Lett.* **15**, 185 (1979).
125. Teng, T. C. *Electron. Lett.* **17**, 627 (1981).
126. Eggermont, G. E. J. & De Groot, J. G. *IEEE Electron Dev. Lett.* **EDL-3**, 156 (1982).
127. Yaron, G. & Hess, L. D. *IEEE Trans. Electron. Dev.* **ED-27**, 573 (1980).
128. Kamins, T. I. & Pianetta, P. A. *IEEE Electron Dev. Lett.* **EDL-1**, 214 (1980).
129. Herbst, D., Bosch, M. A., Lemons, R. A., Tewkesbury, S. K. & Harrison, T. R. *Electron. Lett.* **19**, 12 (1983).
130. Greenwald, A. C. & Little, R. G. *Solid State Technol.* **143** (April, 1979).
131. Leamy, H. J. & Kimerling, L. C. *J. appl. Phys.* **48**, 2795 (1977).
132. Regolini, J. L. et al. *Appl. Phys. Lett.* **34**, 410 (1979).
133. Shibata, K., Inoue, T. & Takigawa, T. *Jap. J. appl. Phys.* **21**, L294 (1982).
134. Yu, T., Soda, J. & Streetman, B. G. *J. appl. Phys.* **51**, 4399 (1980).
135. Knapp, J. A. & Picraux, S. T. *J. appl. Phys.* **53**, 1492 (1982).
136. Golecki, I., Chapman, G. E., Lau, S. S., Tsaur, B. Y. & Mayer, J. W. *Phys. Lett.* **71A**, 267 (1979).
137. Nakata, J. & Kajiyama, K. *Appl. Phys. Lett.* **40**, 686 (1982).
138. Hodgson, R. T., Baglin, J. E. E., Pal, R., Neri, J. M. & Hammer, D. A. *Appl. Phys. Lett.* **37**, 187 (1980).
139. Bomke, H. A., Berkowitz, H. L., Harmatz, M., Kronenberg, S. & Lux, R. *Appl. Phys. Lett.* **33**, 955 (1978).
140. Lue, J. T. *Appl. Phys. Lett.* **36**, 73 (1980).
141. Correria, L. & Pedulli, L. *Appl. Phys. Lett.* **37**, 55 (1980).
142. Nishiyama, K., Ari, M. & Watanabe, N. *Jap. J. appl. Phys.* **19**, L563 (1980).
143. Drowley, C. & Hu, C. *Appl. Phys. Lett.* **38**, 876 (1981).
144. Lue, J. T. & Chao, C. C. *J. appl. Phys.* **53**, 984 (1982).
145. Stultz, T. J., Sturm, J., Gibbons, J. F., Ichiki, S. K. *J. appl. Phys.* **53**, 7109 (1982).
146. Borisenko, V. E. & Labunov, V. A. *Phys. Stat. Sol. A72*, K173 (1982).
147. Vu, D. P., Haond, M., D'Ouville, T. T., Perio, A. & Mingam, H. *Electron. Lett.* **19**, 4 (1983).
148. Tsaur, B.-Y., Donnelly, J. P., Fan, J. C. C. & Geis, M. E. *Appl. Phys. Lett.* **39**, 93 (1981).
149. Fulk, R. T., Russo, C. J., Hanley, P. R. & Kamins, T. I. *Appl. Phys. Lett.* **39**, 604 (1981).
150. Vu, D. P., Haond, M., Bensahel, D. & Dupuy, M. *J. appl. Phys.* **54**, 437 (1983).
151. Geis, M. W., Antoniadis, D. A., Silversmith, D. J., Mountain, R. W. & Smith, H. I. *Appl. Phys. Lett.* **37**, 454 (1980).
152. Fan, J. C. C., Geis, M. W. & Tsaur, B. Y. *Appl. Phys. Lett.* **38**, 365 (1981).
153. Pinizzoto, R. F., Lam, H. W. & Vaandrager, B. L. *Appl. Phys. Lett.* **40**, 388 (1982).
154. Geis, M. W. et al. *Appl. Phys. Lett.* **40**, 158 (1982).
155. Fan, J. C. C., Tsaur, B.-Y., Chapman, R. L. & Geis, M. W. *Appl. Phys. Lett.* **41**, 186 (1982).
156. Hill, C., Butler, A. L. & Daly, J. A. in *Laser and Electron Beam Interactions with Solids*, 238 (Elsevier, Amsterdam, 1982).
157. Young, R. T., Wood, R. F. & Christie, W. H. *Appl. Phys.* **53**, 1178 (1982).
158. Fogarassy, E., Stuck, R., Grob, J. J. & Siffert, P. *J. appl. Phys.* **52**, 1076 (1981).
159. Legge, R. N. & Sharma, K. R. *IEEE Electron. Dev. Lett.* **EDL-1**, 239 (1980).
160. Matsumoto, S., Gibbons, J. F. & Wu, F. C. *J. appl. Phys.* **53**, 7020 (1982).
161. Gibbons, J. F., Lee, K. F., Wu, F. C. & Eggermont, G. E. *J. IEEE Electron Dev. Lett.* **EDL-3**, 191 (1982).
162. Boyd, I. W. thesis, Heriot-Watt University (1982).
163. Csepregi, L., Kennedy, E. F., Mayer, J. W. & Sigmon, T. W. *J. appl. Phys.* **49**, 3906 (1978).
164. Csepregi, L., Mayer, J. W. & Sigmon, T. W. *Appl. Phys. Lett.* **29**, 92 (1976); *Phys. Lett.* **54A**, 157 (1975).
165. Auston, D. H., Golovchenko, J. A. & Venkatesan, T. N. C. *Appl. Phys. Lett.* **34**, 777 (1979).

Evidence for a cooler and drier climate in the Ethiopian uplands towards 2.5 Myr ago

R. Bonnefille

Laboratoire de Géologie du Quaternaire, CNRS, Luminy, 13288 Marseille Cedex 9, France

In east-central Ethiopia (7°N–39°E), at 2,300 m, Pliocene diatomites underlying sediments bearing Acheulian artefacts have provided a rich pollen microflora. The pollen diagram with a Plio-Pleistocene age (2.5–2.35 Myr) is dominated by grasses, Ericaceae and Myrica, whereas the modern pollen rain is dominated by forest taxa. The weak representation of montane forest trees and the abundance of heath moorland and grassland pollen indicate a cooler and drier climate during a time period known to include episodes of global cooling.

RECENT studies from deep-sea cores have led to a considerable improvement in our understanding of palaeoclimatic changes in the late Cenozoic^{1,2}. Although regional variations have been documented, there is evidence for a major interval of cooling between 2.6 and 2.4 Myr ago throughout the Southern Ocean^{3,4} as well as in the Northern Hemisphere⁵. Climatic deterioration was previously shown on land by pollen of the cold praetiglian flora^{6,7} tentatively situated between 2.4 and 2.3 Myr ago⁸. The present study from the uplands of Ethiopia was part of multidisciplinary study of the palaeoenvironments of the prehistoric site of Gadeb^{9–11} and provides a record unique in Africa. It sheds further light on important related issues such as Plio-Pleistocene vegetation history¹² and evolution of early hominids¹³. It brings new information on highland vegetation immediately before the appearance of the stone tool-making hominids and australopithecines in the lowlands.

Climate and phytogeography

In the high plateaus of south-east Ethiopia, at an average elevation of 2,350 m, the Gadeb plain (7°02'–7°13' N, 19°15'–39°28' E) is now drained by the Webi Shebele River. Bounded on the west by the eastern escarpment of the Galla lakes rift, it is limited to the south by the once glaciated¹⁴ Bale Mountains which rise above 4,000 m (Fig. 1). The climatic pattern of the Northern Bale Mountains is typical of that of the central part of the Ethiopian Plateau with a 800–1,000 mm annual rainfall and a wet season from June to September. The highland coniferous areas are characterized by low minimum temperatures. The coldest time of year is also the driest, the main wet season being associated with the convergence of north-east and south-west airstreams¹⁵.

The Gadeb region (2,350 m) lies within the altitudinal range of the forest zone which stands in Ethiopia between 1,800 and 3,300 m (ref. 16). Within the cultivated and grazed 'plain', juniper trees stay as relics of a coniferous forest. In the surrounding areas, two types of forests are still preserved: the Shashemene forest, 50 km to the west, and the Bale Mountains forest, 15 km to the south. At Shashemene, the mixed broad-leaved coniferous forest with *Podocarpus* is a tropical upper montane rain forest¹⁵, known in Ethiopia between 1,800 and 2,500 m under 1,000 mm of mean annual rainfall. Despite a first botanical exploration¹⁷ in 1958–59 and because of difficult access and extent of the massif, the Bale Mountains are not well known. However, recent forest studies have emphasized strong difference in floristic composition between the north and south-east slopes^{15–18}. Altitudinal succession on the northern slope includes montane broad-leaved forest (2,000–2,500 m) followed by a coniferous forest (2,500–3,000 m) with *Juniperus procera* as the dominant element (60–75%) and *Podocarpus gracilior* much less common (4–24%), *Hagenia abyssinica* and *Hypericum revolutum* becoming abundant at the upper limit (3,300 m). The south-facing slopes with a wetter regime show, at the lowest altitudes (1,350–1,600 m), a mixed broad-leaved coniferous forest in which the frequency of the single conifer present, *Podocarpus gracilior*, decreases markedly from south

to north. It is associated with many species among which *Anin-geria adolfriederici*, *Celtis africana*, *Croton macrostachys*, *Polyscias fulva* and *Syzygium guineense* are the most frequent. *Arundinaria alpina* forms a bamboo thicket at elevations over 2,000 m. *Hagenia abyssinica*, *Rapanea melanophloeos* and *Schefflera volkensii* are present in the upper woodlands up to 3,500 m. Areas above 3,500 m are occupied by ericaceous heath thicket and Afro-Alpine grassland with various species of *Alchemilla*, *Labiatae* and *Compositae*¹⁹.

Pollen analyses of modern surface soil samples, taken along an altitudinal transect between 2,000 and 3,800 m in the Bale Mountains, clearly show that, in soil samples, tetrads of *Ericaceae* pollen become abundant only at altitude above 3,500 m (Fig. 2). Relative frequencies of more than 10% are registered in samples collected within the Ericaceous Belt only. Such results are consistent with ericaceous pollen frequencies found in modern pollen studies of other East African mountains: 6–24% on Ruwenzori²⁰ from 3,450 to 4,400 m, 14–16% on Mount Kenya²¹ between 3,300 and 3,500 m, 14–26% on Mount Elgon²² between 3,100 and 3,400 m. Percentages >5% are only recorded in samples collected where heath grow. Pollen of *Ericaceae* is poorly dispersed by air. Its presence in many lake sediments in East Africa, even at low altitude^{23,24}, however, is a result of long distance fluvial transport. In such cases, as well as in mid-altitude lakes²⁵, percentages of *Ericaceae* never exceed 2.5%. Values above 5–10% are reached in lakes

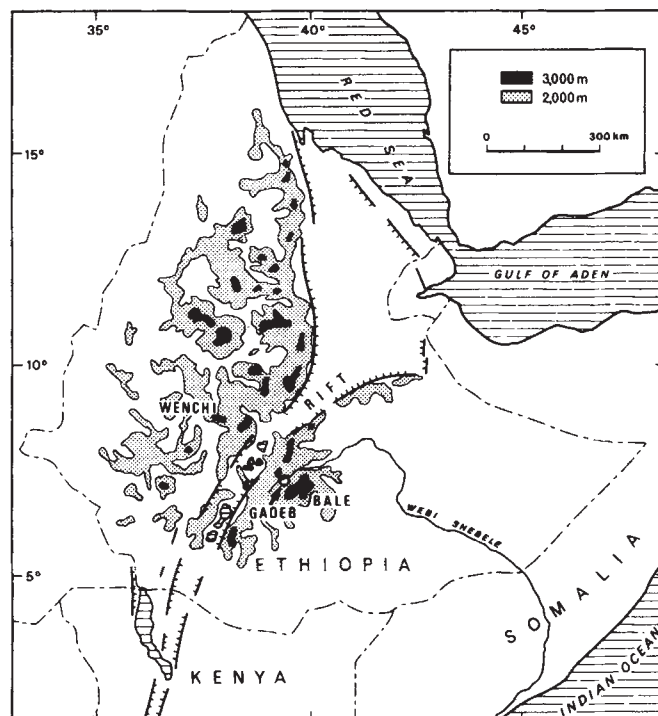


Fig. 1 Locality map.

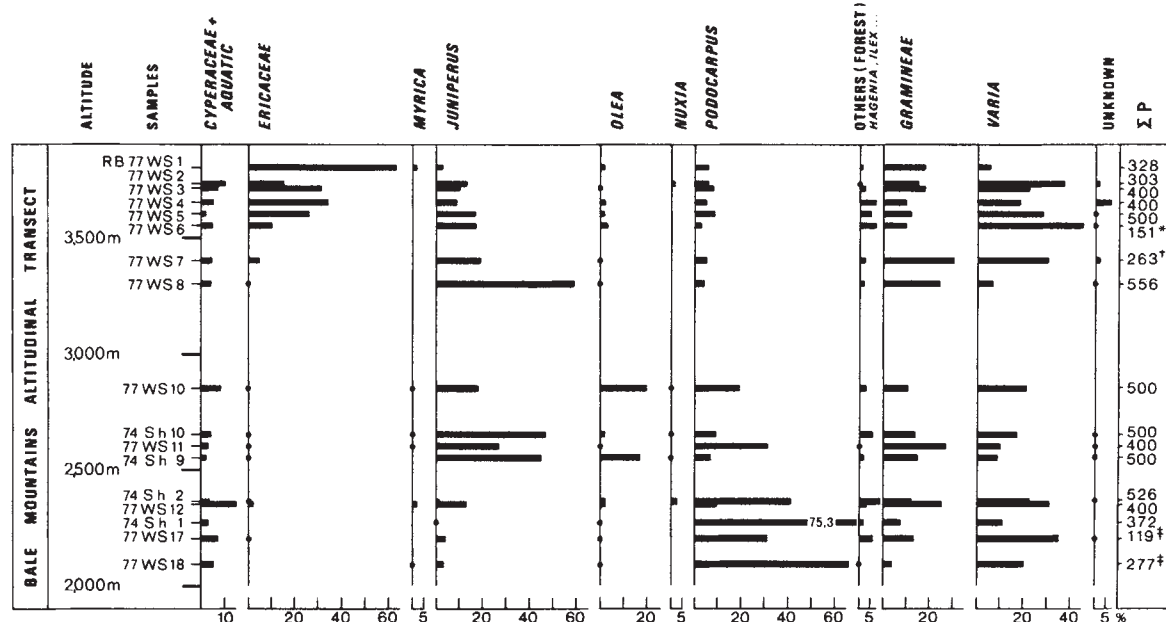


Fig. 2 Distribution of relative pollen frequencies in modern surface samples along an altitudinal transect in the Bale Mountains, % calculated on the total pollen count, over-represented taxa being excluded: **Hypericum*; †*Caryophyllaceae*; ‡*Pteridophyta* (pollen counts by N. and G. Buchet).

situated above 3,000 m, in or near the Ericaceous Belt^{20,21,26,27} (Fig. 3).

Geology

Pliocene lake at Gadeb resulted from the damming of the east-flowing drainage of the ancestral Webi Shebele River by one or more lava flows¹¹ (Fig. 4). Fluvial sands and gravels bearing artefacts overlie diatomites and clays⁹. The oldest diatomites rest on basalts dated to 2.71–2.51 Myr (ref. 11). The fluvial sediments fill erosional channels and postdate a 2.35 ± 0.2 -Myr old 'welded tuff' which conformably overlies the lake deposit¹¹. The upper part of the lacustrine diatomites exposed at Melka Likimi, therefore, dates between 2.51 and 2.35 ± 0.2 Myr (Fig. 4). Preliminary palaeomagnetic measurements suggest that the diatomites were deposited during the Gauss normal epoch, that is before the 2.48 Myr Gauss–Matuyama isochron (M. Williams, personal communication).

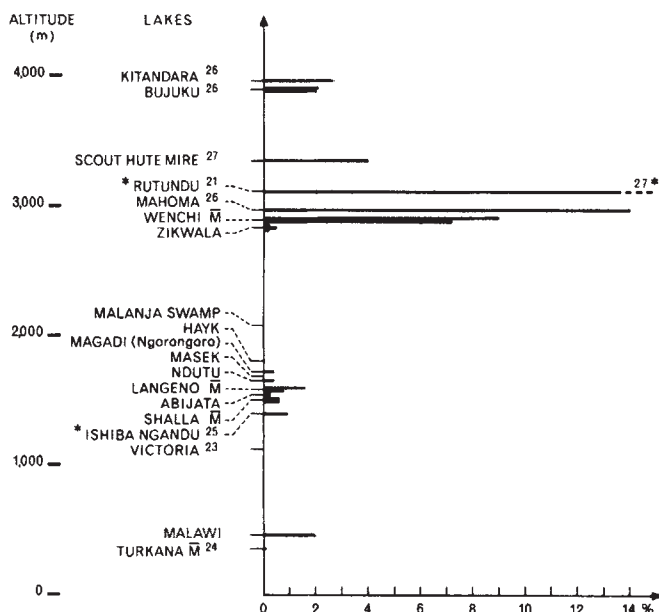


Fig. 3 Relative frequencies of Ericaceae pollen within East African lakes after J. A. Coetzee²¹, A. Hamilton²⁷, D. Hedberg²⁰, R. Kendall²³, D. A. Livingstone^{25,26}, A. Vincens²⁴ and our unpublished data. *% Recalculated.

Pollen analysis

Two hundred and forty samples were collected for palynological studies, from the diatomites and archaeological horizons. The fluvial sediments, volcanic sands and gravels were regarded as unlikely to yield pollen. All the samples from the Oldowan and Acheulian levels (excavations 2E, 8E, 8F) turned out to be sterile. Presented here are the palynological results from the upper part of the lacustrine diatomites.

The sediments were obtained by coring 4.5 m at the base of the Melka Likimi section 6B and sampling at 5–6 cm interval. The same samples were used for diatoms²⁸ and pollen. In the Gadeb diatomite, pollen preservation was found to be exceptionally good. Out of 43 samples, 35 yielded from 10^3 to 10^4 pollen grains g^{-1} . The pollen preservation is better in the lower clayey diatomite than in the upper part of the section 6B where the silty diatomite may have been deposited under highly oxygenated alkaline water as indicated by diatom studies²⁸. Pollen identification has been made by reference to a 6,000-specimen collection of genera and species of the present-day Ethiopian^{16,29–31} and East African flora^{32–34} and present knowledge of pollen morphology available through the literature^{35,36}. We have identified 83 pollen taxa (Table 1), and 5% are still unknown. Considering that several pollen types are grouped by comprehensive taxa such as Umbelliferae, Gramineae, Monocotyledonae, this is a minimum value for the floristic diversity. No pollen taxon was found that corresponded to the 'Tertiary relicts' known in palynological data of the same age in Europe⁸.

The pollen diagram of Fig. 5 summarizes the pollen analyses which will be published in detail elsewhere. It exhibits three important characteristics: a very weak representation of montane forest, an absence of *Podocarpus* and an abundance of Ericaceae followed by an increase of *Myrica*. Three pollen zones can be distinguished on the basis of relative pollen frequencies of *Myrica*, Ericaceae and forest taxa.

(1) Weak representation of montane forest. This is surprising in view of the present position of the Gadeb site within the forest zone. Moreover two tree taxa regularly represented are *Olea* and *Juniperus*, both particularly abundant in dry montane forests. It also contrasts with pollen analyses of modern surface samples along an altitudinal transect between 2,000 and 3,000 m in the Bale Mountains. There, the forest zone is invariably represented by 60–80% arboreal pollen of *Podocarpus*, *Olea*, *Juniperus* and a few others (Fig. 2). Therefore, the forest

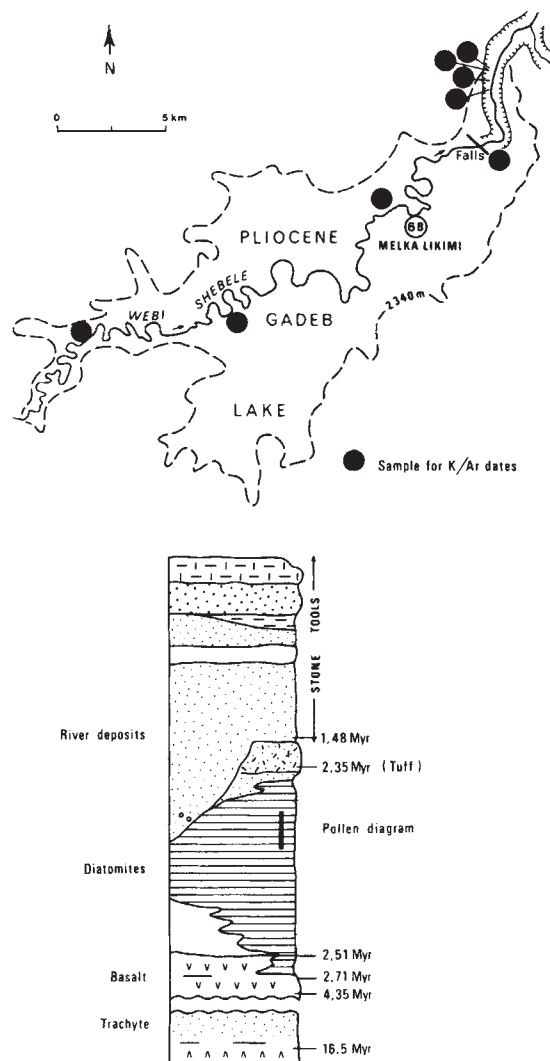


Fig. 4 Summarized stratigraphical sequence and palaeogeographical reconstitution of the Pliocene lake (after M. A. J. Williams *et al.* 1979).

cannot have been present in the immediate vicinity of the Gadeb site while the diatomites were deposited.

(2) **Absence of *Podocarpus*.** *Podocarpus gracilior* is one of the most widespread components of the present-day highland forest of Ethiopia. In south-west Ethiopia, where the Bale Mountains are situated, it is known either alone or in association with other conifers or with Sapotaceae. These associations correspond to an increasing gradient of aridity. The mixed association of Sapotaceae and *Podocarpus* represents the most humid type found under a mean annual rainfall of 1,000–1,400 mm and a dry season lasting from 5 to 6 months. The community with *Juniperus procera* is the driest type in areas with 800–750 mm annual rainfall and a 7 to 8 months dry season. The coniferous forest with both *Podocarpus* and *Juniperus* is intermediate between these two extremes. *Podocarpus gracilior* is absent from forests which are too wet (>1,600 mm) or too dry (<800 mm).

In the Gadeb area today, *Podocarpus gracilior* trees occur in several forests between 2,000 and 2,500 m, at Shashemene on the northern and southern slopes of the Bale Mountains, near Gadeb in the Webi Shebele gorge, along the streams down to 1,700 m near the Galla lakes. *Podocarpus* pollen being well known for its wide dispersal, the record of a single pollen grain on a total pollen count of more than 11,600 grains may reflect a very long distance input from a vegetation several hundred kilometres away. At the time the Gadeb diatomites were deposited, *Podocarpus* was absent from the local and regional vegetation in the catchment of the fossil lake.

Table 1 List of pollen taxa recorded in the Melka Likimi 6B diatomites at Gadeb (Ethiopia)

Acanthaceae	Malvaceae
<i>Acanthus</i> -type	Monocotyledoneae
<i>Hypoestes</i> -type	Myricaceae
<i>Isoglossa</i>	<i>Myrica</i>
<i>Justicia</i>	Myrsinaceae
Anacardiaceae	<i>Myrsine</i> cf. <i>M. africana</i>
<i>Rhus</i> -type	<i>Rapanea</i> -type
Araliaceae	Myrtaceae
<i>Polyscias fulva</i> -type	<i>Syzygium</i>
<i>Schefflera</i> cf. <i>S. abyssinica</i>	<i>Nuxia</i> -type (Loganiaceae)
Balsaminaceae	Oleaceae
<i>Impatiens</i>	<i>Jasminum</i>
Burseraceae	<i>Linociera</i> -type
<i>Commiphora</i>	<i>Olea</i>
Buxaceae	Palmae
<i>Buxus hildebrandtii</i> -type	<i>Borassus/Hyphaene</i>
Campanulaceae	Papilionaceae
<i>Lobelia</i>	<i>Crotalaria</i> -type
Caryophyllaceae	<i>Indigofera</i> -type
Chenopodiaceae	<i>Trifolium</i> -type
<i>Chenopodium</i> -type	undifferentiated
Chenopodiaceae/Amaranthaceae	Plantaginaceae
Combretaceae	<i>Plantago</i>
Commelinaceae	Podocarpaceae
Compositae liguliflorae	<i>Podocarpus</i>
<i>Crepis</i> -type	Polygonaceae
<i>Vernonia</i> -type	<i>Polygonum pulchrum</i> -type
undifferentiated	<i>Rumex</i>
Compositae tubuliflorae	Ranunculaceae
<i>Xanthium</i> -type	<i>Clematis</i>
<i>Artemisia</i>	<i>Ranunculus</i> -type
<i>Carduus</i> -type	<i>Thalictrum</i>
undifferentiated	Rhamnaceae
Cruciferae	Rhizophoraceae
Cupressaceae	<i>Cassipourea</i> -type
<i>Juniperus</i>	Rosaceae
Cyperaceae	<i>Alchemilla</i>
Dipsacaceae	<i>Hagenia</i> cf. <i>H. abyssinica</i>
<i>Dipsacus</i> -type	<i>Prunus</i> cf. <i>P. africana</i>
Ebenaceae	Rubiaceae
<i>Euclea</i>	<i>Anthospermum</i>
Ericaceae	<i>Galium</i> -type
Euphorbiaceae	Salicaceae
<i>Acalypha</i>	<i>Salix</i>
<i>Alchornea</i>	Sapindaceae
<i>Euphorbia</i>	<i>Dodonaea</i> cf. <i>D. viscosa</i>
Geraniaceae	Scrophulariaceae
Gramineae	Sterculiaceae
Guttiferae	<i>Dombeya</i>
<i>Hypericum</i>	Thymelaeaceae
Haloragaceae	<i>Gnidia</i>
<i>Laurembergia</i>	Typhaceae
Hamamelidaceae	<i>Typha</i>
<i>Trichocladus</i>	Umbelliferae
Labiatae	<i>Hydrocotyle</i>
<i>Leonotis</i> -type	undifferentiated
<i>Leucas</i> -type	Urticaceae
<i>Satureja</i> -type	Pteridophyta monolete
undifferentiated (hexacolpate)	Pteridophyta trilete
Linaceae	Anthocerotaceae
<i>Linum</i> -type	

In the fossil pollen diagram (Fig. 5), the two tree pollen taxa which are regularly represented are *Olea* and *Juniperus*, both particularly abundant in dry montane forests. As shown in Fig. 2, *Olea* pollen can be found, though in low frequencies, in modern surface samples above tree line in the Ericaceous Belt where it has been transported by upward winds. *Juniperus procera* is a main component of the Upper Montane Conifer Forest which occupies the highest forest zone between 2,900 and 3,300 m and is able to withstand the frost. In the fossil diagram, its pollen frequencies decrease from the base to the top. When the basal pollen spectrum was deposited, juniper forest may have been present in the catchment of the lake, but it was not very well developed nor close to the lake. In any

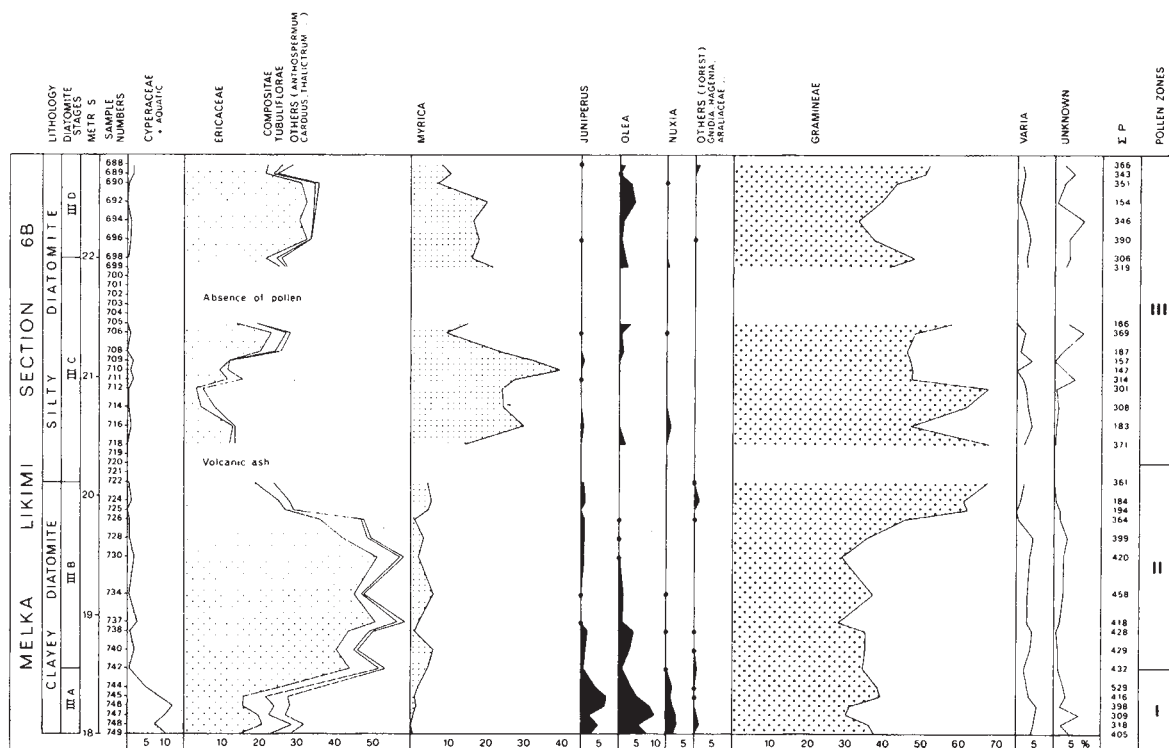


Fig. 5 Pollen diagram of the Melka Likimi diatomites, Upper Webi Shebele basin, Ethiopia.

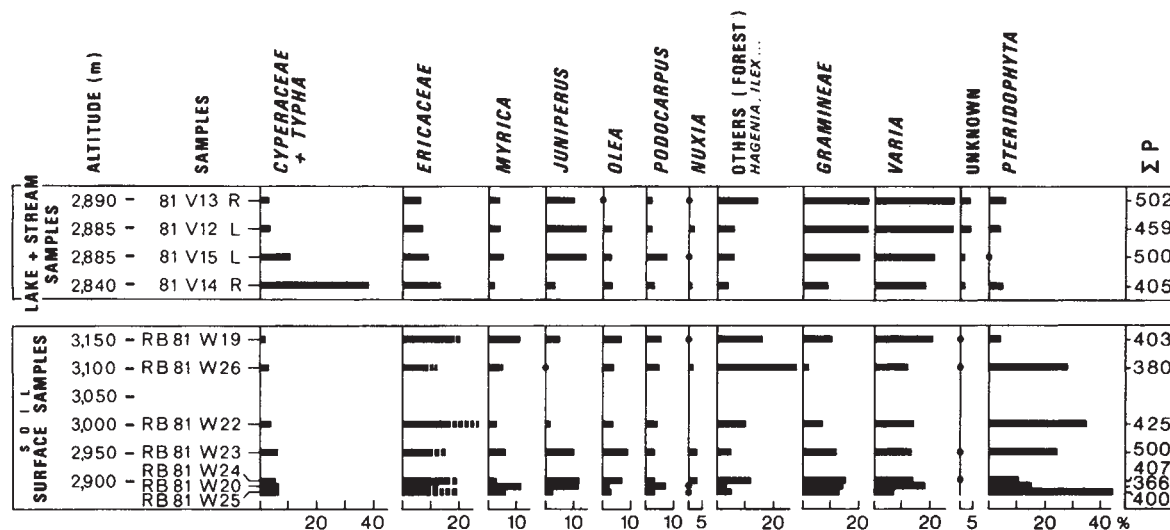


Fig. 6 Distribution of relative pollen frequencies in modern soil and lake samples from the Wenchai crater area, Ethiopia (dotted line represents % calculated with exclusion of Pteridophyta) (pollen counts by G. Buchet).

case, the forest composition, with the absence of *Podocarpus*, attest to a vegetation which nowadays occurs 1,000 m above the altitude of Gadeb, in climatic conditions drier and much cooler than those prevailing around the site today.

(3) **Abundance of Ericaceae.** The pollen diagram from the Melka Likimi 6B diatomites is characterized by the constant presence of Ericaceae tetrads which, in the lower levels, reach relative frequencies of 50%. The preservation of Ericaceae fossil pollen was not good enough to distinguish between the general *Blaeria*, *Erica* and *Philippia*, all of which are represented in Ethiopia by one or more species³⁷, but the fossil pollen resemble *Erica arborea* and *Philippia* most closely. These two tree heaths are the most widespread, forming a dense and continuous shrubby zone called the 'Ericaceous Belt' above the upper limit of the forest on most of the high peaks in East Africa³⁴. The present distribution of Ericaceae in Africa where they are known from north to south and ranging from sea level in temperate localities, to 4,100 m on the East African mountains of the tropical region, raises numerous problems as to their origin and later migration. These ques-

tions have been addressed by earlier writers^{38,39} and again more recently^{40,41}.

Conclusions

Erica arborea is present on all the mountains of Ethiopia. It occurs sporadically above 1,900 m and reaches great importance in the vegetation of the zone between 3,000–3,300 and 3,300–3,700 m in which it becomes the dominant species³⁷. Modern pollen rain studies have shown that pollen frequencies over 50% are recorded only in surface samples from the Ericaceous Belt. This applies not only to the Bale Mountains in Ethiopia (Fig. 2) but also to other mountains in East Africa where it was discussed by A. Hamilton²⁷. The Ericaceae pollen curve of the Gadeb diagram points to dense Ericaceae vegetation in the immediate vicinity of the site, whether this implies that the site was in the Ericaceous Belt proper or that we have sampled a more local ericaceous community, on the edge of the lake, within the forest belt.

Despite the fact that Ericaceae pollen tetrads are known to be locally dispersed^{20,22}, published data from lake sediments

illustrated in Fig. 3 have shown that a small amount is transported widely by streams and rivers. Relative values above 10%, however, are reached only in the lakes situated around or above 3,000 m, within the Ericaceous Belt.

Parallel sampling of surface soil and lake sediments was possible in the area of the Wench crater lake, near 3,000 m on the northeastern Ethiopian Plateau. The results are summarized in Fig. 6 and they clearly show that Ericaceae pollen frequencies are not increased in lake sediments by the catchment of the lake, lower values being registered in the lake sediments and not in the surface samples of the same area.

Considering again the fossil data of the Gadeb diatomites, it can be seen that in addition to the high Ericaceae percentages, the forest taxa never exceed 15%, a value which can be reached in modern surface samples from the Ericaceous Belt²¹ (Fig. 2) and which is due to upward dispersal by air currents above the tree line. As pointed out before, the most abundant forest taxa are the moderately dispersed *Olea* and the frost-resistant upper dry mountain forest *Juniperus*. The fossil pollen diagram can logically be interpreted as sampling a vegetation found today in the Ericaceous Belt proper. It could be argued that abundance of Ericaceae could be a result of the pioneer character of heath on acid soils or burnt and disturbed ground. Volcanic activity during the diatomite deposition is illustrated by a volcanic ash layer (level 720) which seems to have had very little influence on the Ericaceae pollen curve (Fig. 5). On the basis of K/Ar dates, the 30 m diatomite sequence at Gadeb may have been deposited during 200,000 yr maximum. Thus, the pollen diagram of Fig. 5 reflects a vegetation which may have persisted during 30,000 yr at the most. In any case, the time period of development of the Ericaceae maximum is too long to explain as being the result of the pioneer character of the heath vegetation.

From the arguments expressed above it can be inferred, that the diatomite pollen diagram represents the Ericaceous Belt proper. Tectonic changes cannot have been the cause of the occurrence of the fossil pollen spectra at this low altitude as the only tectonic activity which has affected the area since the Pliocene was the uplift of the Gadeb basin. In East Africa, the lower limit of the Ericaceous Belt lies at about 2,600–3,400 m, and the upper limit fluctuates between 3,550 and 4,100 m. Thus, the existence of the Ericaceous Belt 2.5-Myr ago at an altitude at least 1,000 m below its present lower limit, implies climatic conditions much different from the present.

The fluctuation of *Myrica*, one of the main pollen components, appearing after the Ericaceae maximum, is not fully understood. Its increase, after the Ericaceae decline, may reflect variations of the local vegetation around the lake. *Myrica* is a readily recognizable pollen on the generic level. Of the four

species known from East Africa, only two are widely distributed but their pollen is very similar. *Myrica kandiana* is a swamp plant, *Myrica salicifolia* is a dryland demanding drought-tolerant species. In some surface samples from East Africa that we have studied, *Myrica* appears to be a component of the modern pollen rain in the Ericaceous Belt. However, we do not know of any modern example showing *Myrica* pollen frequencies as high as those in the Gadeb pollen diagram. The fossil data suggest the presence of xerothermic peatbog vegetation postulated by previous authors³⁹ to explain the discontinuous geographical distribution of Myricaceae and Ericaceae on the East African peaks. A peat-bog of this nature must have existed sometime between 2.51 and 2.35 Myr ago. The highest frequencies of *Myrica*, associated with great percentages of Gramineae and very few *Juniperus* and *Olea* makes it likely that *M. salicifolia* was the species present at Gadeb. In more recent pollen diagrams from East Africa, relatively abundant *Myrica* during the glacial maximum before 12,000 yr BP is attributed by A. Hamilton to the widespread occurrence of dry montane woodland under a generally rather arid climate²⁷.

In tropical regions today, the Ericaceous Belt is concentrated in cool high mountains, the temperature is low, at least during the night. The aridity is never severe as the atmospheric moisture keeps the general humidity high⁴¹ but plants have to withstand drought from strong insolation during the day. In the Ericaceous Belt, on most mountains, the climate is usually wet during most of the year, but it experiences a dry season, and consequently the vegetation has a mesoxerophytic temperate character³⁴. The temperate character of the fossil flora is shown also by the occurrence of temperate genera such as *Crepis*, *Prunus*, *Rhamnus*, *Thalictrum*⁴², Dipsacaceae, *Juniperus*^{43,44} or boreal genera with northern temperate affinity such as *Carduus*⁴². The xerophytic character of the flora can be inferred from the presence of Araliaceae, *Hypericum*, *Olea*, *Anthospermum*⁴⁵.

The pollen diagram from Gadeb shows that 2.5–2.35-Myr ago on Ethiopian highlands, the climate must have been colder and in strong contrast with a substantial dry season. But more information is needed to assess climatic significance to the *Myrica* and Ericaceae fluctuations within this time period.

The palynological studies are part of a wider study of Ethiopian palaeoanthropology directed by Professor J. C. Clark of the University of California, Berkeley. I thank F. Gasse for providing samples, J. R. Flenley, O. Hedberg, D. Livingstone, N. Morner, M. Rossignol-strick and E. M. Van Zinderen Bakker for critical comments. D. Livingstone and M. Williams for improving the final version of the manuscript. The work was done with technical help from G. Riollot and G. and N. Buchet. It has been financed by the CNRS and the NSF grant to J. D. Clark.

Received 17 February; accepted 24 March 1983.

1. Frakes, L. A. *Climates throughout Geologic Time* (Elsevier, Amsterdam, 1979).
2. Pearson, R. *Climates and Evolution* (Academic, London, 1978).
3. Hays, J. D. *Am. geophys. Un. Antarct. Res. Ser. S*, 125–184 (1965).
4. Weaver, F. H. *Florida. State. Univ. Sediment. Res. Lab. Rep.* 36, 142 (1973).
5. Shackleton, N. J. *Nature* 270, 216–219 (1977).
6. Tsukada, M. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 3, 49–64 (1967).
7. Leopold, E. in *Aspect of Palynology* (eds Tschudy, R. H. & Scott, R. A.) 377–438 (Wiley, New York, 1969).
8. Zagwijn, W. H. in *Ice Ages, Ancient and Modern* (eds Wright, A. E. & Moseley, F.) 135–152 (Seel House, Liverpool, 1975).
9. Clark, J. D. & Kurashina, H. *Nature* 282, 33–39 (1979).
10. Clark, J. D. et al. in *Palaeoecology of Africa* Vol. 11 (ed. Van Zinderen Bakker, E. M.) 145–147 (Balkema, Rotterdam, 1979).
11. Williams, M. A. J., Williams, F. M., Gasse, F., Curtis, G. H. & Adamson, D. A. *Nature* 281, 29–33 (1979).
12. Axelrod, D. I. & Raven, P. H. in *Biogeography and Ecology of Southern Africa* (ed. Werger, M. J. A.) 77–130 (Junk, The Hague, 1978).
13. Cronin, J. E., Boaz, N. T., Stringer, C. B. & Rak, Y. *Nature* 298, 113–122 (1981).
14. Messerli, B., Winiger, M. & Rognon, P. in *The Sahara and Nile* (eds Williams, M. A. H. & Faure, H.) 87–132 (Balkema, Rotterdam, 1980).
15. Chaffey, D. R. *South-west Ethiopia Forest Inventory Project* (Land Resources Development Centre, Surbiton 1979).
16. Logan, W. E. M. *An Introduction to the Forest of Central and Southern Ethiopia* (Imperial Forestry Institute, Oxford, 1946).
17. Mooney, H. F. *Proc. Linnean Soc. Lond.* 172, Pt. 2, 126–152 (1963).
18. Henry, P. W. T. *South-west Ethiopia Forest Inventory Project* (Land Resources Development Centre, Surbiton, 1973).
19. Hedberg, O. *Boissiera* 24, 71–74 (1975).
20. Hedberg, O. *Oikos* 5, 137–165 (1954).

21. Coetzee, J. A. in *Palaeoecology of Africa* Vol. 3 (ed. Van Zinderen Bakker, E. M.) 1–100 (Balkema, Cape Town, 1967).
22. Hamilton, A. C. & Perrott, R. A. *Pollen et Spores* 22, 437–468 (1980).
23. Kendall, R. L. *Ecol. Monogr.* 39, 121–176 (1969).
24. Vincens, A. thesis, Univ. Aix-Marseille II (1982).
25. Livingstone, D. A. *Limnol. Oceanogr.* 16, 349–356 (1971).
26. Livingstone, D. A. *Ecol. Monogr.* 37, 25–52 (1967).
27. Hamilton, A. C. *Environmental History of East Africa* (Academic, London, 1982).
28. Gasse, F. *Les diatomées lacustres plio-pléistocènes du Gadeb* (Revue algologique, Paris, 1980).
29. Pichi-Sermolli, R. *Webbia* 13, 15–132 (1957).
30. Cufodontis, G. *Enumeratio plantarum aethiopiae spermatophyta* (Jardin botanique National Belgique, Bruxelles, 1974).
31. Richard, A. *Tentamen Flora Abyssinicae* (Inst. system. Bot. Uppsala, 1982).
32. Turill, W. B. & Milne-Redhead, M. A. (eds) *Flora of Tropical East Africa* (Crown Agents, London, 1952–82).
33. Hedberg, O. *Symb. Bot. upsal.* 15, 1–409 (1957).
34. Hedberg, O. *Svensk. bot. Tidskr.* 45, 140–202 (1951).
35. Bonnefille, R. *Pollen et Spores* 22, 15–72 (1971).
36. Bonnefille, R. & Riollot, G. *Pollens des Savanes d'Afrique Orientale* (CNRS, Paris, 1980).
37. Pichi-Sermolli, R. & Heiniger, H. *Webbia* 9, 9–48 (1953).
38. Chevalier, A. *Bull. Soc. bot. Fr.* 70, 855–870 (1923).
39. Christ, H. *Arch. Sci. phys. nat. Genève*, 3rd period 28, 369–376 (1892).
40. Schnell, R. *Introduction à la phytogéographie des pays tropicaux. IV La flore et la végétation de l'Afrique tropicale*, part 2 (Gauthier-Villars, Paris, 1977).
41. Van Zinderen Bakker, E. M. *Argumenta geogr.* 12, 89–97 (1977).
42. Good, D. O. *The Geography of the Flowering Plants* (Longman, London, 1974).
43. Hedberg, O. *Webbia* 19, 519–529 (1965).
44. Bader, F. J. W. *Webbia* 19, 531–544 (1965).
45. Lebrun, J. *Bull. Jard. bot. Brux.* 30, 258–388 (1960).

Sr, O, and Pb isotope evidence for origin and evolution of Etive Igneous Complex, Scotland

J. A. P. Clayburn*, R. S. Harmon†, R. J. Pankhurst‡ & J. F. Brown*

* Department of Geology and Mineralogy, University of Oxford, Oxford OX1 3PR, UK

† Department of Geological Sciences, Southern Methodist University, Dallas, Texas 75275, USA

‡ British Antarctic Survey, c/o Institute of Geological Sciences, 64 Gray's Inn Road, London WC1X 8NG, UK

A multi-isotope approach has been applied to the origin of one of the largest and latest Newer Granite complexes of the Scottish Caledonides. Contributions from both mantle and unexposed late Proterozoic lower continental crust are recognized, indicating that granite magmatism is a mechanism for both primary crustal growth and reworking of older crustal rocks.

CURRENT discussion concerning the development of the continental crust centres on the extent to which it has been progressively formed throughout geological time by direct differential chemical segregation from the mantle versus an initial mantle derivation early in Earth history followed by the episodic recycling of pre-existing crustal material. Critical to this problem is an understanding of the petrogenesis of Phanerozoic

granitoids because granitic material is a major residual component of old, stabilized continental crust. Present geochemical and isotopic evidence have been taken to support either case¹⁻³. For example, Moorbath⁴ has presented Pb-Sr isotopic evidence that a mantle component is present in granitic rocks of all ages, a feature which can be considered indicative of continuous primary continental accretion. By contrast, Allègre and co-

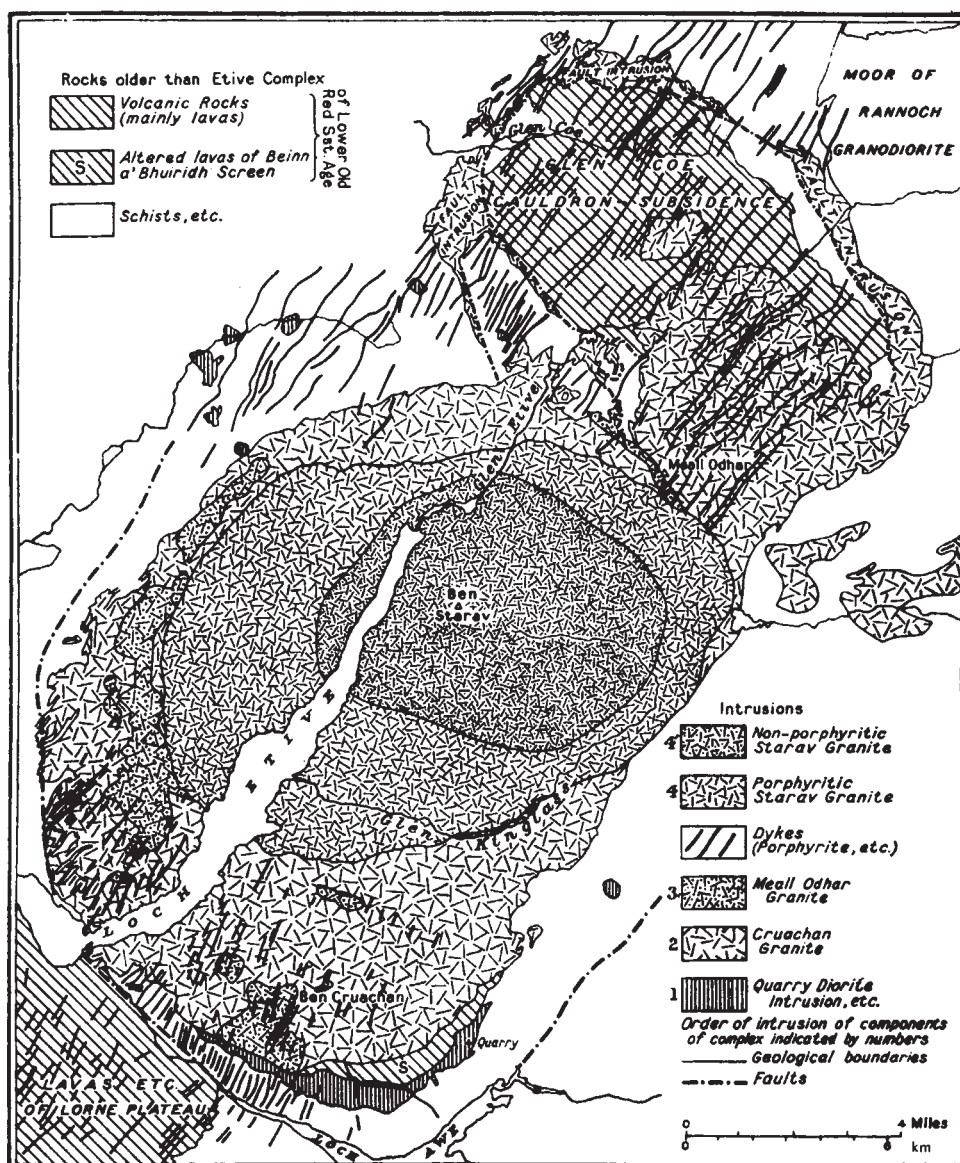


Fig. 1 Geological map of the Etive Igneous Complex, western Grampian Highlands, Scotland (from ref. 44). Shown are the major granitic intrusive phases of the complex and their relationship to adjacent Dalradian meta-sedimentary country rock and other Caledonian granitoids, associated volcanic rocks of equivalent age, and the Glen Gairn subsidence cauldron. Note that, in the north-east, the Cruachan Granite emerges concordantly into the Fault Intrusion, which is not shown as a separate stratigraphical unit.

workers^{5,6} argue from Nd-Sr isotopic relationships that most Phanerozoic granitoids appear to contain a large crustal component, thus favouring the idea of extensive recycling of continental crust in granitoid generation.

The Caledonian Orogenic Belt developed along the margin of the North American-Scandinavian shield as the result of subduction and plate collision associated with the closure of the Iapetus Ocean in late Precambrian to early Ordovician times. In Britain the emplacement of an extensive granitoid province both accompanied and followed the deformational and metamorphic events of the Caledonian Orogeny. We have undertaken a multi-isotopic study of one important Caledonian granitoid complex in the western Grampian Highlands of Scotland in order to assess the role of the mantle versus that of continental crust in the generation of Phanerozoic granitoids.

Etive Igneous Complex

The Etive Igneous Complex is a polyphase, calc-alkaline intrusion which outcrops over an area of ~300 km² around Loch Etive in the western Grampian Highlands of Scotland (Fig. 1) and constitutes one of the largest of the post-tectonic (late) Caledonian granitoid complexes in Britain⁷⁻⁹. The plutonic complex was passively emplaced into down-faulted late Proterozoic (Dalradian) metasediments in early Devonian time. In age and origin there is a clear association between the Etive plutonic complex and the outcrops of Old Red Sandstone (ORS) calc-alkaline volcanic rocks which form the Lorne Plateau to the south-west and Glen Coe subsidence cauldron to the north-east (Fig. 1).

The earliest intrusive phase of the Etive Complex is the Quarry Intrusion (G1), a small lens of diorite/quartz-diorite which outcrops at the southeastern margin of the complex between the altered Beinn a Bhuraidh andesite screen and the local Dalradian metasedimentary country rock. The Cruachan Intrusion (G2) is the oldest major phase of the Etive complex. It is compositionally heterogeneous, ranging from a fine-grained

tonalite in the southern lobe, where xenoliths of Dalradian quartzite and schist are common, to a coarse-grained biotite-hornblende granodiorite in the northern lobe. In the north-east the Cruachan Intrusion merges concordantly into the finer-grained Fault Intrusion (G3), a heterogeneous series of magmas injected along the Glen Coe ring fault. The next phase of the complex to be emplaced was the Meall Ohdar Granite (G4), a pink K-feldspar and quartz granite which occurs only within the outcrop of the Cruachan intrusion as a series of sheet and dyke intrusions. Penecontemporaneous with intrusion of the Meall Ohdar Granite was the injection of a suite of north-east-trending Porphyrite Dykes (G5) which cut across the early phases of the plutonic complex. The centre of the Etive complex is composed of two concentric granites, the outer Porphyritic Starav Granite (G6) and the inner leucocratic Central Starav Granite (G7).

Previous isotopic reconnaissance studies on the Etive Igneous Complex have reached petrogenetic conclusions which are difficult to reconcile. Low initial Sr-isotopic ratios (<0.706) (ref. 10) and the lack of a substantial 'inherited' zircon component¹¹ for G2 have been taken to indicate derivation from the mantle or a young lower crustal source, whereas Nd-Sr isotope relationships (G2: $\epsilon_{Nd} = -4.9$, $\epsilon_{Sr} = +11$; G6: $\epsilon_{Nd} = -9.5$ to -9.9 ; $\epsilon_{Sr} = +14$ to $+20$) have been cited¹² as evidence for a substantial component of old continental crust to the Etive magmas. More recently, Frost and O'Nions¹³ have suggested that both mantle and Dalradian-type upper crust were important in the origin and evolution of the Etive magmas.

Table 1 Whole-rock geochronological data for the Etive Igneous Complex and associated Old Red Sandstone volcanics

Sample	Rb (p.p.m.)	Sr (p.p.m.)	⁸⁷ Rb/ ⁸⁶ Sr	⁸⁷ Sr/ ⁸⁶ Sr
Lorne Lavas (pre-G1) (andesites and felsites)				
A190	55	582	0.272	0.70634
A071	78	250	0.900	0.70957
A072	49	911	0.153	0.70560
G01	71	883	0.231	0.70621
G02	122	100	3.516	0.72506
G03	110	135	2.343	0.71808
D23	72	1,180	0.179	0.70560
D45	67	140	1.378	0.71258
D73	180	86	6.028	0.73893
Meall Ohdar granite (G4)				
A106	103	86	3.514	0.72565
A107	115	104	3.201	0.72414
A108	130	35	10.63	0.76673
A190	157	37	12.30	0.77667
A192	99	38	7.46	0.76796
A197	115	42	7.83	0.75030
A198	110	13	25.16	0.84732
Central Starav granite (G7)				
A045	158	128	3.568	0.72530
A046	179	111	4.654	0.73142
A047	178	108	4.782	0.73297
A204	178	204	2.513	0.71962
A205	177	243	2.113	0.71747

Sample details and localities are given in refs 14, 15. Rb/Sr ratios were determined by X-ray fluorescence spectrometry variously in Oxford and London, UK with errors assumed to be $\pm 1\%$, except for sample A198 where a very low Sr content yields an error of $\pm 2\%$. Errors on ⁸⁷Sr/⁸⁶Sr ratios, also determined partly in Oxford and partly in London, are $\pm 0.01\%$.

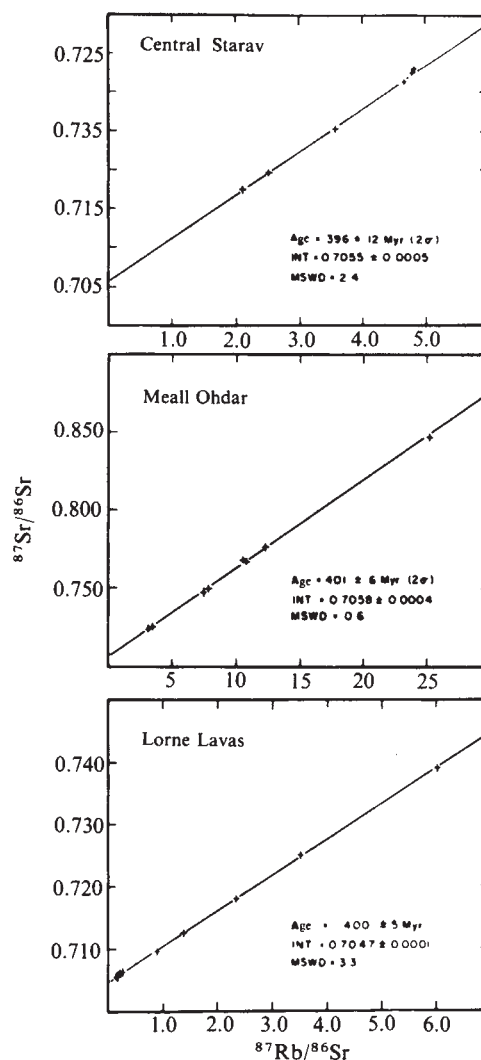


Fig. 2 Rb-Sr whole-rock isochrons for the Lorne Lavas, Meall Ohdar Granite, and Central Starav Granite. Thirlwall (personal communication) has analysed samples of the Lorne lavas which are less altered than these and has obtained an age of ~410 Myr.

Results

Rb-Sr geochronological data are given in Table 1 and isochrons for the G4 and G7 phases of the Etive Complex as well as the temporally associated, pre-G1 Lorne Lavas presented in Fig. 2. Figure 2 shows that the entire Etive Igneous Complex was emplaced over a very short interval at ~400 Myr into down-faulted Upper Proterozoic (Dalradian) metasediments and penecontemporaneous stratified Lorne and Glen Coe volcanics (Fig. 1).

Unpublished Rb-Sr mineral data support the whole-rock (WR) chronology summarized above, and indicate that the mineral phases of each intrusive phase were essentially in Sr-isotopic equilibrium at the time of crystallization. Brown¹⁴ gives biotite-WR ages of 410 ± 7 Myr for G1, an average of 410 Myr for three G2 samples, and 414 ± 10 Myr for G6. A discordantly young age of 373 ± 6 Myr for G3 presumably reflects subsequent open system behaviour. Clayburn¹⁵ reports combined whole-rock and mixed feldspar isochrons of 389 ± 4 and 385 ± 6 Myr for two G7 samples used in the present study, again possibly indicating slight resetting compared with the whole-rock isochron of Fig. 2.

Major and trace element data from previous work^{8,14,16} and this study (Table 2) indicate that the Etive Complex has calc-alkaline affinities, and is similar in composition to other large volume, passively-emplaced, post-tectonic Caledonian intrusions in Scotland¹⁷⁻¹⁹. Most notable are the high Ni, Cr, Sr, and Ba contents of the early G1 and G2 phases as well as the high K/Na, K/Rb, and Rb/Sr ratios and strong light rare-earth elements (LREE) enrichment observed for the complex as a whole. Differentiation indices for the Etive Complex increase in the order G1-G2-G3, G5-G6-G7-G4. Comparison with the compositional variations of the Lorne and Glen Coe lavas was previously used⁸ to argue that the Etive trend was due to magmatic differentiation by simple liquid line of descent fractional crystallization. The major and trace element data

presented here (Fig. 3) define smooth curvilinear trends which are interpreted to result from a complex magmatic history involving limited crystal fractionation within individual intrusive phases coupled with episodic magma mixing in a deep magma chamber. The reversal in the general fractionation trend after G4 is thought to represent the influx of a new batch of magma into the Etive magma chamber. The high contents of Ni (>100 p.p.m.) and Cr (>150 p.p.m.) in the least evolved of the G1 samples are similar to those observed in the most primitive of the genetically related ORS lavas which Thirlwall²⁰ considered to be indicative of a derivation from a mantle source.

Overall ranges in the five isotopic parameters of interest for the Etive Complex are: $(^{87}\text{Sr}/^{86}\text{Sr})_i = 0.7044-0.7058$, $\delta^{18}\text{O}_{\text{SMOW}} = 7.2-9.8\%$, $^{206}\text{Pb}/^{204}\text{Pb} = 16.714-17.185$, $^{207}\text{Pb}/^{204}\text{Pb} = 15.314-15.447$, and $^{208}\text{Pb}/^{204}\text{Pb} = 36.371-36.911$ (Table 2), making the Etive Complex the least evolved isotopically of the post-tectonic Caledonian granitoids of the Northern and Grampian Highlands. There is a sympathetic increase in both $(^{87}\text{Sr}/^{86}\text{Sr})_i$ and $^{18}\text{O}/^{16}\text{O}$ ratios, as well as decrease in all three Pb-isotope ratios, with increasing SiO_2 and Rb/Sr within the Etive Complex (Table 2, Fig. 3). The slight drop in $(^{87}\text{Sr}/^{86}\text{Sr})_i$ from 0.7058 to 0.7051 and $\delta^{18}\text{O}$ from 8.9 to 8.8% between G4 and G5 is consistent with the idea of the injection of a new batch of magma into the Etive Chamber at this time. Because radiogenic isotopic composition of a magma is unaffected by fractional crystallization, and the stable isotopic composition of a magma is only slightly sensitive at plutonic temperatures ($\geq 1.5\%$)²¹, the systematic changes we observe in Sr-, O-, and Pb- isotopic composition within the Etive Complex must reflect either primary variations in a heterogeneous source region which was partially melted to produce the Etive magmas, or a variable degree of contamination of partial melts derived from a homogeneous source before or during emplacement.

Figure 4 shows plots of $\delta^{18}\text{O}-(^{87}\text{Sr}/^{86}\text{Sr})_i$ and $^{206}\text{Pb}/^{204}\text{Pb}-(^{87}\text{Sr}/^{86}\text{Sr})_i$ for the Etive Igneous Complex. Also shown are our

Table 2 Summary of age, chemical, and isotopic relationships within the Etive Igneous Complex

Intrusive phase*	Age (Myr)	SiO ₂ (wt %)	K ₂ O (wt %)	Sr (p.p.m.)	Rb/Sr	K/Na	K/Rb	K/Ba	$\delta^{18}\text{O}_{\text{SMOW}}$ (‰)	$(^{87}\text{Sr}/^{86}\text{Sr})_i$	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
G7 n	402 ± 10	4	4	5	5	4	3	3	5	5	3	4	3
r		74-76	4.73-5.00	108-243	0.73-2.8	1.38-1.72	234-254	165-217	8.9-9.8	0.7055±	16.714-16.860	15.314-15.318	36.276-26.295
m		75.6	4.86	159	1.5	1.55	243	187	9.4		16.798	15.316	36.286
G6 n	(400)	3	3	5	5	3	3	3	3	3	3	3	2
r		69-72	4.39-4.56	296-739	0.16-0.43	1.05-1.53	296-360	32-39	9.3-9.7	0.70476-0.70542	16.753-16.765	15.322-15.363	36.266-36.403
m		70.9	4.47	491	0.26	1.34	327	34.7	9.5	0.70507	16.759	15.343	36.335
G5 n	(400)	4	4	6	6	4	4	4	3	3			
r		63-70	3.61-4.55	381-811	0.07-0.27	0.94-1.37	286-365	22-30	8.6-9.2	0.70480-0.70517			
m		66.4	3.96	598	0.17	1.10	324	25.9	8.8	0.70501			
G4 n	400 ± 4	3	3	8	8	3	3	3	6				
r		75-77	5.14-5.34	13-104	1.1-8.5	1.29-1.52	328-443	98-318	8.3-9.8	0.7058±			
m		76.2	5.23	56	3.1	1.44	404	173	9.0				
G3 n	(400)	4	4	4	4	4	4	4		1	1	1	1
r		54-68	2.01-3.80	723-1,266	0.3-0.10	0.59-0.86	337-399	12-23		0.70532	16.88	15.395	36.304
m		60.4	2.88	1,078	0.07	0.76	365	17.5		0.70532	16.822	15.395	36.304
G2 n	(400)	12	12	15	15	12	11	11	6	13	2	2	2
r		56-70	2.72-4.67	420-1,306	0.05-0.31	0.64-1.45	473-289	8-40	7.2-8.7	0.70474-0.70550	16.925-17.178	15.337-15.345	36.466-36.798
m		61.8	3.65	822	0.15	1.08	368	20.5	8.2	0.70493	17.052	15.341	36.632
G1 n	(400)	5	5	5	5	5	5	5	1	1	1	1	1
r		56-62	1.89-3.21	943-1,235	0.03-0.07	0.87-1.23	392-416	19-23	7.2	0.70445	17.185	15.447	36.911
m		59.7	2.62	1,070	0.06	1.02	400	20.1	7.2	0.70445	17.185	15.447	36.911
Lorne Lavas	400 ± 5									0.7047±			

* Given for each intrusive phase are the number of samples analysed (n), together with the range (r) and mean value (m) for each parameter.

† All Pb-isotope data are for feldspar separates with age corrections for $^{206}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ ratios mostly less than 0.02 and 0.005 respectively.

‡ From Fig. 2.

§ Provides an older limit for the age of the Etive Complex, but see Fig. 2 legend.

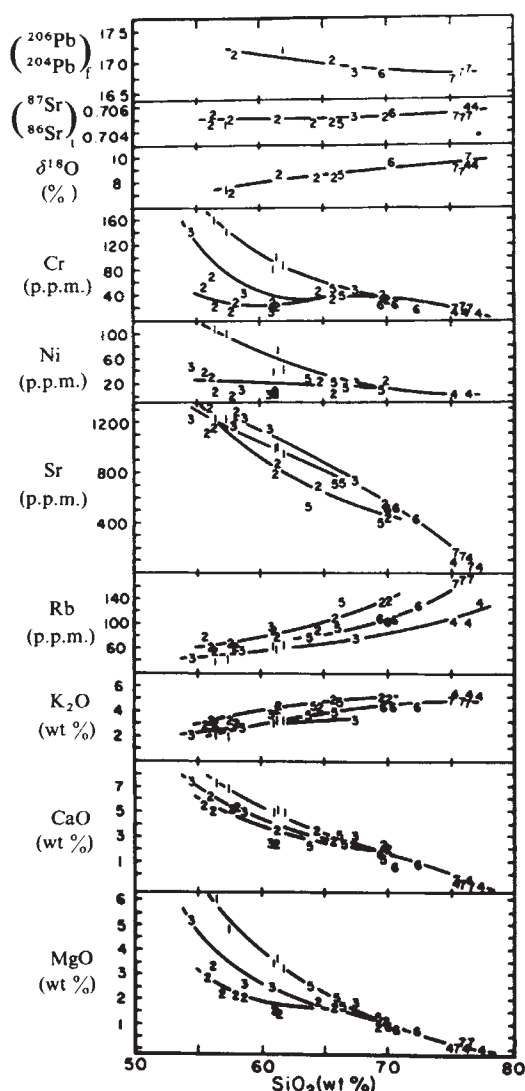


Fig. 3 Harker variation diagrams for selected major elements for samples from the Etive Igneous Complex. The numbers refer to the seven intrusive phases described in the text.

determinations or best estimates of the stable and radiogenic isotopic composition of possible source regions from which the Etive magmas might have been derived together with various crustal reservoirs with which the magmas might have interacted before final emplacement and crystallization. Isotopic data for the Strathspey-Laggan Complex¹⁴, a post-tectonic, forcefully-emplaced Caledonian granitoid also in the Western Highlands, and the post-tectonic granitoids of the Southern Uplands (Doon, Criffell and Fleet)²² are shown for comparison. The main question of concern here is whether the isotope systematics of the Etive Complex reflect derivation from a single source or represent a mixture between isotopically distinct reservoirs indicative of post-derivation contamination.

Figure 3 shows that there is a clear tendency for the Etive magmas to become enriched in ^{18}O , ^{87}Sr , but depleted in radiogenic Pb from the more mafic G1 diorite (57–61% SiO_2) and G2 tonalite (56–61% SiO_2), through the G2 and G3 granodiorites, to the G4, G6, and G7 granites (71–77% SiO_2). Clearly this coherent behaviour amongst the three isotope systems argues strongly for a common origin for the observed isotopic variations. Lithospheric reservoirs which might be envisaged to have a role in the origin and evolution of the Etive magmas include the upper mantle, subducted or underplated Iapetus oceanic crust, and stabilized Precambrian lower and/or upper crust of variable age between 2,900 and 600 Myr. Two important points are clear from the plots of the isotope data in Fig. 4. First, it is obvious that the Etive magmas were not

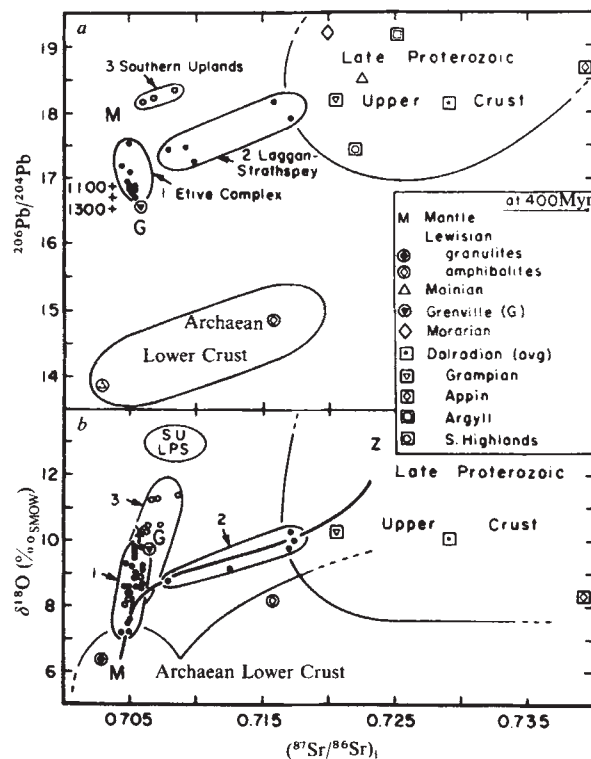


Fig. 4 Relationships between $^{206}\text{Pb}/^{204}\text{Pb}$ and $(^{87}\text{Sr}/^{86}\text{Sr})_i$ ratios (a) and $^{18}\text{O}/^{16}\text{O}$ – $(^{87}\text{Sr}/^{86}\text{Sr})_i$ ratios (b) for the Etive Igneous Complex (●). Also shown for comparison are Pb–Sr and O–Sr isotope relationships for 2, the Laggan–Strathspey Complex (○) considered by Clayburn¹⁵ to be derived by anatexis of Dalradian metasediments, and for 3, the granitoid intrusions (Doon, Criffell, and Fleet) of the Southern Uplands (○) considered by Halliday *et al.*²² to have resulted from mixing of mantle derived melts with those derived by *in situ* partial melting of the local Lower Palaeozoic geosynclinal sediments (SULPS) in which the intrusions are located. Also shown for reference are the Pb–Sr and O–Sr isotopic compositions at 400 Myr of source regions from which the Etive magmas might have been derived together with the upper crustal materials with which the Etive magmas might have interacted before final emplacement and crystallization. The numbered crosses in a represent the Sr–Pb isotope composition of sub-crustal mantle at 1,100, 1,200, and 1,300 Myr. The O-isotope datum for Grenville rocks is an average for rocks of intermediate composition from ref.36. See text for discussion.

derived from a single homogeneous source with the variations observed in their chemical composition subsequently produced solely by magmatic differentiation before emplacement. The O-, Sr-, and Pb-isotope data presented in Fig. 4 demonstrates that an origin solely from heterogeneous Archaean (Lewisian) lower crust or late Proterozoic (Dalradian) upper crust is unlikely. In both the O–Sr and Sr–Pb isotope systems the Etive data define sub-linear arrays of points which trend away from contemporary mantle towards the field estimated for middle Proterozoic lower crust. The Nd-isotope data of Hamilton *et al.*¹² and Frost and O'Nions¹³ for the Etive Complex are consistent with this conclusion, and also rule out a derivation from MORB or a MORB-type, depleted mantle source. Second, the isotopic trends displayed by the Etive magmas indicate that both the local Upper Proterozoic (Dalradian) metasedimentary upper crust in which the Etive Complex now resides and the upper continental crust through which the Etive magmas passed before emplacement at their present structural level, have played an insignificant part in determining the Sr-, O-, and Pb-isotopic composition of the intrusion. This is especially well illustrated in Fig. 4a. Recent theoretical modelling^{23–25} has shown that source contamination (such as large-ion lithophile enrichment during subduction) and magma mixing are both two-component mixing processes which define linear or hyperbolic mixing curves on stable-radiogenic isotope diagrams,

whereas combined fractional crystallization–assimilation is a three-component process which results in S-shaped curves which for much of their evolutionary path may not directly extrapolate to the source or contaminant composition. The curves shown in Fig. 4b for contamination of a mantle-derived melt by mixing with middle Proterozoic lower crust (M–X) and fractional crystallization of a mantle-derived melt concomitant with assimilation of late Proterozoic upper crust (M–Z) rule out the latter process in the evolution of the Etive magmas.

Note that the isotopic behaviour exhibited by the Etive Complex contrasts sharply with that for the Strathspey–Laggan Complex, considered to be derived by crustal anatexis which resulted from decompression of the Dalradian metamorphic pile due to rapid post-deformation uplift and erosion¹⁴. In Fig. 4 the Strathspey–Laggan data plot on both isotope diagrams as linear arrays which trend away from an unradiogenic lower crustal source region towards the field for radiogenic upper crust, corresponding closely to the hypothetical curve for combined fractional crystallization–assimilation shown on the $\delta^{18}\text{O}$ –($^{87}\text{Sr}/^{86}\text{Sr}$)_i plot (Fig. 4b). Such clear evidence for a late Proterozoic or younger upper crustal component in the granitic rocks of the Etive complex is completely lacking. Rather, the origin of the Etive magmas seems best explained in terms of contamination of a mantle-derived melt by moderately unradiogenic middle Proterozoic lower crust or by anatexis of heterogeneous lower crust containing a juvenile mantle-derived component.

Several authors^{15,26–28} have presented evidence that the Lewisian lower crust is not sufficiently radiogenic to be the predominant source of the Pb and ^{87}Sr in the British Caledonian granitoids. Likewise, amphibolite facies gneisses, similar in isotopic composition to Lewisian amphibolites cannot be directly involved in the genesis of the Etive magmas. These rocks are characterized by high Th/U ratios which have resulted from decoupling of Th from U at the amphibolite grade of metamorphism²⁹, and so show disparity in plots including ^{208}Pb .

Given that the Lewisian is not a viable major component in the petrogenesis of the Etive magmas, the identity of the required unradiogenic lower crustal component is problematic. Three younger Precambrian orogenic events have been identified in Scotland which might reasonably be expected to have contributed to the lower crust beneath the Grampian Highlands. These are the ~1,900–1,600 Myr Laxfordian event, the ~1,100 Myr Grenville event, and the ~800 Myr Morarian event^{14,28,30–33}. Both the Laxfordian and Morarian events principally involved the reworking and reconstitution of older Archaean crust with little introduction of juvenile Pb added to the crust. This contrasts strongly with the Grenville event which is recognized as a major period of new crustal addition in other regions. Isotope data for relevant Grenville rocks are extremely limited. Figure 4 plots fields based on the best estimates of Sr-, O-, and Pb-isotope composition at 400 Myr for Grenville province rocks in the Canadian shield^{34–36}. In Scotland, only the Ardour gneiss is recognized as a possible Grenville-age crustal addition, but with a mean $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.708 at the time of formation at 1,200 Myr (Brook *et al.*³⁷) and Rb/Sr ratios of 0.7–1.5, this would have had quite radiogenic Sr during the Caledonian Orogeny, and is probably not typical of lower crustal correlatives. We know of only a few Pb-isotope measurements in Grenville rocks³⁸ so we also estimated the composition in 1,100 Myr sub-crustal mantle from the well-known ore-Pb growth curves^{39,40} and the $^{206}\text{Pb}/^{204}\text{Pb}$ ratio of 16.90 is plotted in Fig. 4a. This falls well within the range of ratios measured for the Etive feldspars. Clearly the low-melting point, felsic to intermediate component of the Grenville might well be a suitable endmember to provide the lower crustal component required for the Etive magmas, provided that as a new crustal segment it was subjected to high-grade metamorphism and some U-depletion soon after accretion. It would probably not be a suitable single source for the Etive magmas, however, because there are severe theoretical difficulties in the large-scale bulk melting of continental crust to produce large volumes of

dioritic or tonalitic magma⁴¹.

There is, however, geophysical evidence to suggest that some Lewisian basement extends beneath the Grampian Highlands⁴². This would imply that the sub-Grampian lower crust under central Scotland is a complex basement which is in part juvenile Grenville granulite, in part Grenville contaminated by Lewisian at ~1,100 Myr, and in part pristine Lewisian granulite. If this is so, then the actual lower crustal component to the Etive magmas would probably lie in a position in Fig. 4 intermediate between the Grenville and Lewisian fields.

Finally, note that previously published Sr- and Nd-isotopic data for the Etive Igneous Complex are consistent with the interpretation presented here. The Sr-isotopic data of Halliday *et al.*¹⁰ and Hamilton *et al.*¹² fall within the range that we have observed. Also, the Sr–Nd isotope data and t_{CHUR} ages presented by Hamilton *et al.*¹² for three Etive samples recognize a crustal component to the Etive magmas which had a previous crustal residence time of at least 400–800 Myr, consistent with a Grenville component. Intermediate to felsic Grenville crust extracted from CHUR at 1,200 Myr would evolve to produce ϵ_{Nd} values and Sm/Nd ratios similar to those observed in the late Etive magmas by 400 Myr. That the $^{143}\text{Nd}/^{144}\text{Nd}$ ratios in the G6 granites are lower (0.51162–0.51164) than in the G2 granodiorite (0.51187) reflects a greater proportion of lower crustal component in the late phases of the intrusion.

If our interpretation of the data in Fig. 4a is correct, then those Caledonian granitoids plotting towards the upper right side on an O–Sr diagram (Strathspey and so on) could have been derived by mixing of melts derived from the postulated Grenville-type basement with melts from the Dalradian-type upper crust. These appear on all isotopic systems to lack the juvenile mantle-type component present in the Etive magmas and other large post-tectonic Caledonian batholiths.

Conclusions

The Sr-, O-, and Pb-isotopic data presented here suggest that the granitic rocks of the Etive Igneous Complex were produced as a result of melting of both contemporary mantle and older lower crust. We envisage a juvenile melt derived from an enriched mantle source overlying the subducted Iapetus slab to initiate melting within the lower crust, and subsequently mix with this melt during prolonged residence in a lower crustal magma reservoir. The resultant very dry, hybrid magma was then intruded rapidly by fracture-assisted emplacement into the upper crust. Each successive magma pulse intruded through earlier lavas or granitic rock was derived from the same source and, as a result, suffered only minor contamination by local Dalradian metasedimentary upper crust. The progressive depletion of successive intrusive phases in radiogenic Pb, as well as enrichment in ^{87}Sr and ^{18}O , indicates that the G4 to G7 phases of the intrusion contain a larger lower crustal component than the earlier G1 to G3 phases. This implies that lower crustal melting continued throughout the duration of igneous activity at the Etive Complex and increasingly dominated the isotopic character of the magma chamber with time. This process is seen as one of combined primary crustal growth coupled with assimilation of older, stabilized continental crust.

Analytical support was provided by NERC studentships to J.A.P.C. and J.F.B. and grants to Oxford University, the Scottish Universities Research and Reactor Centre, and the Institute of Geological Sciences. We thank J. Borthwick for technical assistance in the laboratory and J. R. O'Neil for O-isotope analyses. Figure 1 is reproduced with permission of the director, Institute of Geological Sciences, Crown copyright.

Received 23 September 1982; accepted 14 April 1983.

1. Armstrong, R. L. *Rev. Geophys.* **6**, 175–199 (1968).
2. Brown, G. C. in *Origin of Granite Batholiths—Geochemical Evidence* (eds Atherton, M. P. & Tarney, J.) 106–115 (Shiva, Orpington, 1979).
3. Fyfe, W. S. *Chem. Geol.* **23**, 89–114 (1978); *Phil. Trans. R. Soc. A* **291**, 433–445 (1979).
4. Moorbath, S. *Chem. Geol.* **20**, 151–187 (1977); *Phil. Trans. R. Soc. A* **288**, 401–413 (1978).
5. Allegre, C. J., Ben Othman, D., Polve, M. & Richard, P. *Phys. Earth planet. Inter.* **19**, 293–306 (1979).
6. Allegre, C. J. & Ben Othman, D. *Nature* **286**, 335–342 (1980).
7. Clough, C. T., Maufe, H. B. & Bailey, E. B. *Q. J. geol. Soc. Lond.* **65**, 611–676 (1909).

8. Bailey, E. B. & Maufe, H. B. *Mem. Geol. Sur. U.K.* (1960).
9. Pankhurst, R. J. & Sutherland, D. S. in *Igneous Rocks of the British Isles* (ed. Sutherland, D. S.) 149–190 (Wiley, Chichester, 1982).
10. Halliday, A. N., Aftalion, M., van Breemen, O. & Jocelyn, J. *Spec. Publ. geol. Soc. Lond.* **8**, 653–661 (1979).
11. Pidgeon, R. T. & Aftalion, M. *Geol. Soc. Spec. Iss.* **10**, 183–220 (1978).
12. Hamilton, P. J., O'Nions, R. K. & Pankhurst, R. J. *Nature* **287**, 279–284 (1980).
13. Frost, C. S. & O'Nions, R. K. *EOS* **63**, 461 (1982).
14. Brown, J. F. thesis, Oxford Univ. (1975).
15. Clayburn, J. A. P. thesis, Oxford Univ. (1981).
16. Plant, J. A., Brown, G. C., Simpson, P. R. & Smith, R. T., *Trans. Instn Min. Metall.* **B89**, 198–201 (1980).
17. Nockolds, S. R. & Mitchell, R. L. *Trans. R. Soc. Edinb.* **61**, 533–575 (1946).
18. Mercy, E. L. P. in *The British Caledonides* (eds Johnson, M. R. W. & Stewart, F. H.) 189–215 (Oliver and Boyd, Edinburgh, 1963).
19. Halsam, H. W. *J. Petrol.* **9**, 84–104 (1968).
20. Thirlwall, M. F. *Earth planet. Sci. Lett.* **58**, 27–50 (1982).
21. Chivas, A. R., Andrew, A. S., Sinha, A. K. & O'Neil, J. R. *Nature* **300**, 139–143 (1982).
22. Halliday, A. N., Harmon, R. S. & Stephens, W. E. *J. geol. Soc. Lond.* **137**, 329–348 (1981).
23. Taylor, H. P. *Earth planet. Sci. Lett.* **47**, 243–254 (1979).
24. DePaolo, D. J. *Earth planet. Sci. Lett.* **53**, 189–202 (1981).
25. James, D. E. *Ann. Rev. Earth planet. Sci.* **9**, 311–344 (1981).
26. Blaxland, A. B., Aftalion, M. & van Breemen, O. *Scott. J. Geol.* **15**, 139–151 (1979).
27. Chapman, H. J. & Moorbath, S. *Nature* **268**, 41–42 (1977).
28. van Breemen, O., Halliday, A. N., Johnson, M. R. W. & Bowes, D. R. *Geol. J. Spec. Iss.* **10**, 81–106 (1978).
29. Chapman, H. J. thesis, Oxford Univ. (1978).
30. Moorbath, S. & Park, R. G. *Scott. J. Geol.* **8**, 51–74 (1971).
31. Bickerman, M., Bowes, D. R. & van Breemen, O. *J. geol. Soc. Lond.* **131**, 237–254 (1975).
32. Brook, M., Brewer, M. S. & Powell, D. *Nature* **260**, 515–517 (1976).
33. Pankhurst, R. J. & Pidgeon, R. T. *Earth planet. Sci. Lett.* **31**, 55–58 (1976).
34. Krogh, T. E. & Hurley, P. M. *J. geophys. Res.* **73**, 7107–7125 (1968).
35. Fullagar, P. D., Hatcher, R. D. & Merschat, C. E. *Southeastern Geol.* **20**, 69–77 (1969).
36. Shieh, Y.-N. & Schwarcz, H. P. *Can. J. Earth Sci.* **15**, 1773–1782 (1978).
37. Brook, M., Brewer, M. S. & Powell, D. *Nature* **260**, 515–517 (1976).
38. Fletcher, I. R. & Farquhar, R. M. *Can. J. Earth Sci.* **19**, 239–253 (1982).
39. Stacey, J. S. & Kramers, J. D. *Earth planet. Sci. Lett.* **26**, 207–221 (1975).
40. Cumming, G. L. & Richards, J. R. *Earth planet. Sci. Lett.* **28**, 155–171 (1975).
41. Wyllie, P. J. in *Energetics of Geological Processes* (eds Saxena, S. K. & Bhattacharji, S.) 389–433 (Springer, New York, 1977).
42. Durrance, E. M. *Geol. Mag.* **113**, 251–261 (1976).
43. Westbrook, G. K. & Borradaile, G. J. *Scott. J. Geol.* **14**, 213–224 (1978).
44. Johnstone, G. S. *British Regional Geology: The Grampian Highlands* (HMSO, London, 1966).

A cellular oncogene (c-Ki-ras) is amplified, overexpressed, and located within karyotypic abnormalities in mouse adrenocortical tumour cells

Manfred Schwab, Kari Alitalo, Harold E. Varmus & J. Michael Bishop

Department of Microbiology and Immunology, University of California Medical Center, San Francisco, California 94143, USA

Donna George

Department of Human Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA

The cellular oncogene c-Ki-ras is amplified 30- to 60-fold in cells of the mouse adrenocortical tumour Y1. The amplified oncogene is located in double minute chromosomes and in a homogeneously staining chromosomal region, common karyotypical anomalies of tumour cells. The amounts of c-Ki-ras specific mRNA and of the protein (p21) encoded by the amplified gene are correspondingly elevated. Amplification and enhanced expression of cellular oncogenes may contribute to the genesis and/or maintenance of at least some naturally occurring tumours.

VERTEBRATE genomes contain a group of 'cellular oncogenes' whose enhanced expression or alteration by mutation may be involved in the genesis of naturally occurring tumours¹. The possibility that cellular oncogenes are involved in tumorigenesis is supported by four lines of evidence. First, transduction of cellular oncogenes into the genomes of retroviruses can endow the viruses with oncogenic potential¹. Second, tumorigenesis by retroviruses that lack oncogenes may begin when the integration of viral DNA enhances the expression of a nearby cellular oncogene^{2–6}. Third, DNAs from a variety of tumours contain genes that can transform rodent cells to a neoplastic phenotype^{7,8}. In several cases, the responsible genes have proven to be previously identified cellular oncogenes^{9–14}, and at least one of these is apparently a mutant version of the wild-type allele^{12–14}. Fourth, several chromosomal translocations that typify certain human and mouse tumours relocate known cellular oncogenes and may affect either the structure or the activity of the oncogenes^{15–22}. Similarly, translocation of a small domain of DNA has activated *c-mos* in a mouse plasmacytoma²³.

Gene amplification can vastly increase the expression of individual genetic loci²⁴. Two karyotypic abnormalities signal the presence of amplified genes: double minute chromosomes (DMs) and homogeneously staining regions of chromosomes (HSRs)^{25,26}. DMs are small, usually paired chromatin bodies that lack centromeres and thus segregate unpredictably at cell division, whereas HSRs are domains within chromosomes

where the characteristic banding pattern has been replaced by chromatin material devoid of morphological detail. Although the origins and nature of DMs and HSRs first emerged from the study of gene amplification in cells resistant to cytotoxic drugs^{27,28}, both of these karyotypic abnormalities have also been observed in a large number and variety of tumour cells that have not been exposed to chemotherapeutic agents²⁵.

In pursuit of the possibility that enhanced expression of cellular oncogenes may contribute to the genesis of naturally occurring tumours, we have deliberately examined tumour cells containing DMs and/or HSRs in the hope of finding amplification of cellular oncogenes. Here we report that the cellular oncogene *c-Ki-ras* is amplified 30- to 60-fold in the Y1 cell line established from the mouse adrenocortical tumour LAF₁, that the amplified oncogene resides on DMs and HSRs in the tumour cells, and that both polyadenylated RNA transcribed from *c-Ki-ras* and the protein encoded by the gene (p21^{c-Ki-ras}) are exceptionally abundant in the tumour cells. Enhanced expression of cellular oncogenes as an apparent consequence of gene amplification has been observed recently in at least two other settings—the HL-60 line of human promyelocytic leukaemia cells^{29,30} and the COLO 320 line of human neuroendocrine tumour cells³¹. It therefore seems reasonable to suspect that amplification and enhanced expression of cellular oncogenes may contribute to the genesis and/or maintenance of at least some naturally occurring tumours.

c-Ki-ras expression in Y1 cells

Screening of cell lines for expression of cellular oncogenes was carried out using the assay illustrated in Fig. 1. Polyadenylated RNA from tumour cells was transcribed into radioactive cDNA with reverse transcriptase. The cDNA was then hybridized to nitrocellulose filters bearing grids of molecularly cloned DNA representing 13 of the known cellular oncogenes. In tests of RNA from the Y1 cell line, only c-Ki-ras was represented in amounts appreciably above control values provided by RNA from NIH 3T3 cells (Fig. 1). Y1 cells are available as two sublines, one containing DMs, the other containing HSRs³². Both sublines showed enhanced expression of c-Ki-ras.

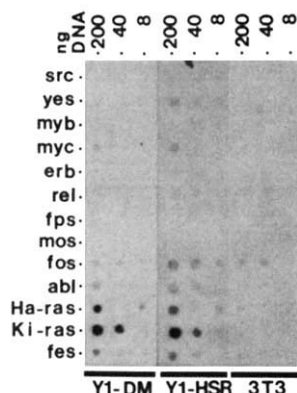


Fig. 1 Analysis of expression of cellular oncogenes in Y1-DM, Y1-HSR and NIH 3T3 cells. Oncogene-specific inserts were excised with restriction endonucleases from plasmid or bacteriophage clones of avian (*src*, *yes*, *myb*, *myc*, *erb*, *rel*, *fps*), murine (*mos*, *fos*, *abl*, *Ha-ras*, *Ki-ras*), or feline (*fes*) retroviruses and purified by electrophoresis in agarose gel. DNA was serially diluted to 100 ng, 20 ng and 4 ng μl^{-1} and heat-denatured. Aliquots (2 μl) corresponding to 200 ng, 40 ng and 8 ng were spotted onto dry nitrocellulose filters (Sartorius) presoaked in 0.3 M sodium citrate, 3 M sodium chloride ($20\times\text{SSC}$). The filters were dried and baked for 2 h at 80°C . The Y1-HSR cells were originally obtained from G. Sato at the University of California at San Diego, Y1-DM cells were from American Type Culture Collection. Cells were grown in Eagle's minimal essential medium supplemented with 10% fetal bovine serum. For isolation of polyadenylated RNA, $\sim 10^8$ cells were lysed with 0.5% SDS in 0.1 M NaCl, 20 mM EDTA, 50 mM Tris-HCl, pH 7.6 (lysis buffer). The cellular proteins were hydrolysed with $500\text{ }\mu\text{g ml}^{-1}$ proteinase K (Merck) for 30 min at 37°C . The DNA was sheared with a Virtis homogenizer. The NaCl concentration was raised to 0.5 M and polyadenylated RNA was bound to oligo(dT)-cellulose (Collaborative Research) during a period of 4 h under agitation. The oligo(dT)-cellulose was centrifuged (1,500 r.p.m., 3 min) at room temperature, the supernatant discarded, and the oligo(dT) was resuspended in lysis buffer. This step was carried out three times. The oligo(dT) was then poured into a column, and the polyadenylated RNA was eluted with 10 mM Tris, pH 7.6, 1 mM EDTA. The RNA was precipitated in the presence of 0.3 M sodium acetate, pH 8.0, with 3 volumes of ethanol at -70°C . Radioactive cDNA was prepared with the reverse transcriptase of avian myeloblastosis virus (AMV) using oligomers of calf thymus DNA as random primers and [^{32}P]dCTP (specific activity 800 Ci mmol^{-1}) to achieve specific activities of $2\text{--}5\times 10^7$ c.p.m. per μg template. Hybridizations were performed in Denhardt's mixture (0.02% each of bovine serum albumin, Ficoll and polyvinylpyrrolidone), $3\times\text{SSC}$, 40% formamide, 50 mM HEPES, pH 7.2, $200\text{ }\mu\text{g ml}^{-1}$ salmon sperm DNA, $150\text{ }\mu\text{g ml}^{-1}$ yeast tRNA, for 24 h at 41°C . Approximately 10^7 c.p.m. were applied to each filter in 15 ml hybridization mixture. Filters were washed in $1\times\text{SSC}$, 0.1% SDS at 30°C and autoradiographed. Filters were re-used for up to six times without detectable loss of signal. For denaturation of the oncogene-cDNA hybrids, the filters were soaked with slight agitation in 0.2 M NaOH, 0.1% SDS at room temperature for 10 min, rinsed with tapwater, neutralized for 3 min in 0.5 M Tris-HCl, pH 7.0, and washed in 0.1% SDS. Filters were dried at room temperature on filter paper, preincubated in hybridization mixture and used as described above.

Fractionation of RNA from Y1 cells by electrophoresis in agarose gels and hybridization of the RNA to cDNA for Ki-ras revealed three species of RNA with sizes of 5.2 kilobases (kb), 2.0 kb and 1.2 kb (Fig. 2). The 5.2-kb and 2.0-kb RNAs have been observed previously in NIH 3T3 fibroblasts and the 416B line of mouse haematopoietic precursor cells³³. We used RNA from both of these cell lines as controls. The expression of c-Ki-ras is enhanced approximately 25-fold in the 416B cells³³. As a consequence, RNAs transcribed from c-Ki-ras in these cells were readily detectable, in quantities roughly equivalent to those found in Y1-HSR cells (Fig. 2). The 416B cells contained the 5.2-, 2.0- and 1.2-kb RNAs found also in Y1 cells (the 1.2-kb species is not readily visible in the figure). By contrast, barely detectable amounts of only the 5.2-kb and 2.0-kb RNAs were apparent in the NIH 3T3 cells (Fig. 2). We cannot yet account for the novel 1.2-kb RNA, and we cannot discern whether it is present in the NIH 3T3 cells. It is unlikely that it is the transcript of c-Ha-ras, an oncogene related to c-Ki-ras. First, the Ha- and Ki-ras oncogenes do not cross-react appreciably in the hybridization conditions used for the experiment in Fig. 2. Second, when the RNA species were annealed with probe for the Ha-ras gene in conditions of reduced stringency, only weak signals were obtained for each of the three species, with ratios of intensity similar to those observed with the Ki-ras probe (not shown). From the data illustrated in Fig. 2 and other, similar analyses, we can again estimate that the expression of c-Ki-ras is enhanced approximately 60-fold in the Y1-DM cells and about 25-fold in the Y1-HSR cells.

Normal cells contain small amounts of a protein encoded by c-Ki-ras that has a molecular weight of 21,000 ($p21^{\text{c-Ki-ras}}$). We asked whether the amount of $p21^{\text{c-Ki-ras}}$ in Y1 cells is elevated. Cells were metabolically labelled; $p21^{\text{c-Ki-ras}}$ was then precipitated with monoclonal antibody³⁴ and analysed by gel electrophoresis. Extracts from 416B cells were analysed in parallel; in addition, the mouse myeloma cell line NS-1 was tested for comparison. Figure 3 shows that both the Y1-DM and the Y1-HSR cells, as well as the control 416B cells, contain large amounts of $p21^{\text{c-Ki-ras}}$, by contrast, NS-1 cells show the anticipated low level of $p21^{\text{c-Ki-ras}}$.

We used partial hydrolysis with the V8 protease of *Staphylococcus aureus* to substantiate the identity of the putative $p21^{\text{c-Ki-ras}}$ recovered from Y1 cells. The 21,000-molecular weight (M_r) proteins obtained from 416B and Y1 cells yielded identical polypeptide products on proteolysis (data not shown). An additional protein with M_r of $\sim 16,000$ was specifically precipitated from extracts of the Y1-DM cells. Mapping with V8 protease demonstrated that this protein is related to

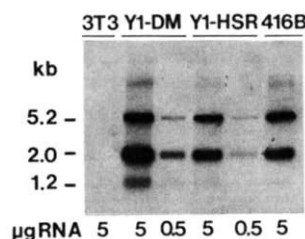


Fig. 2 Characterization of RNA transcribed from c-Ki-ras in the polyadenylated RNA of mouse Y1-DM and Y1-HSR cells. NIH 3T3 and 416 B cells were used as controls. RNA was isolated as described in Fig. 1 legend and dissolved in sample buffer containing 20 mM MOPS (pH 7.0), 1 mM EDTA, 5 mM sodium acetate, 50% (v/v) formamide and 2.2 M formaldehyde. Samples containing amounts of RNA as indicated were heated for 10 min at 60°C and electrophoresed in 0.8% agarose gels. Electrophoresis buffer was the same as the sample buffer. The position of ribosomal RNA was visualized by ethidium bromide staining, and the RNA was transferred to nitrocellulose papers in $20\times\text{SSC}$. As molecular probe for detection of c-Ki-ras, we used the *SsrII-EcoRI* fragment isolated from clone HiHi-3 of v-Ki-ras³⁷. Conditions for hybridization were as described in Fig. 1 legend, except that 50% formamide was used.

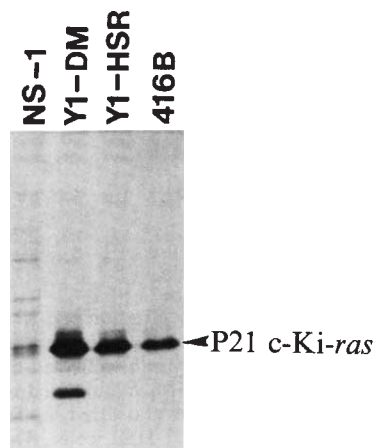


Fig. 3 Immunoprecipitation of p21^{c-Ki-ras} from the Y1-DM and Y1-HSR cells. Mouse myeloma NS-1 and the 416B haematopoietic precursor cells were used as controls. Cells were metabolically labelled with ³⁵S-methionine (200 µCi ml⁻¹, 500 Ci mmol⁻¹; NEN) for 3 h. Lysates were prepared and immunoprecipitated by the monoclonal antibody Y13-259 (given by E. Scolnick). The p21 immunoprecipitated with this antibody from cells transformed with Ha-MSV, v-Ha-ras or c-Ha-ras has a different mobility during gel electrophoresis than p21 from cells transformed with Ki-MSV; the p21 detected in 416B cells is of the type found in Ki-MSV-transformed cells³⁴. The immunoprecipitates were dissolved in electrophoresis sample buffer and electrophoresed in a 15% SDS-polyacrylamide gel. The gel was dried and subjected to fluorography.

p21^{c-Ki-ras} (data not shown) but we cannot account for the origin of this previously undescribed form of Ki-ras protein.

We used immunofluorescence microscopy to examine the location of p21^{c-Ki-ras} in Y1 cells. A strong fluorescent signal was obtained when the monoclonal antibody was used to stain Y1 cells (data not shown). The pattern of fluorescence suggested that much of the protein is associated with the plasma membrane, as shown previously for the related protein encoded by the *ras* gene of Harvey sarcoma virus³⁵. By contrast, little or no specific staining was seen in NIH 3T3 cells, which contain small quantities of c-Ki-ras apparently below the level of detection of immunofluorescence. From data of the sort illustrated in Fig. 3, we estimated that the amount of labelled p21^{c-Ki-ras} in Y1-DM cells is reproducibly higher than that in the Y1-HSR cells, and that both forms of Y1 cells contain more p21^{c-Ki-ras} than do 416B cells. These estimates are in accord with the findings for c-Ki-ras RNA described above.

c-Ki-ras amplification

Given the presence of DMs and HSRs in Y1 cells, it seemed likely that the enhanced expression of *c-Ki-ras* might be due to amplification of the cellular oncogene. We therefore analysed the quantity and topography of *c-Ki-ras* in the DNA of Y cells by hydrolysis with restriction endonucleases. The signals obtained in the experiment illustrated by Fig. 4 indicated that *c-Ki-ras* is amplified approximately 60-fold in Y1-DM cells and ~30-fold in Y1-HSR cells. These deductions are in accord with the magnitude of the enhancement of expression of *c-Ki-ras* described above, and with the results of previous studies using randomly chosen molecular clones representing the DNA of DMs in Y1-DM cells³⁶. Cleavage of DNAs from Y1-DM, Y1-HSR and NIH 3T3 cells with *EcoRI* produced fragments of 9.5 kilobase pairs (kbp), 8.5 kbp, 1.5 kbp and 0.5 kbp that hybridized with *Ki-ras* probe (Fig. 4). We have also analysed DNA from the liver and testes of an LAF₁ mouse, the strain from which the Y1 line was derived; the results were identical to those seen with DNA from NIH 3T3 cells (data not shown). The 9.5-, 1.5- and 0.5-kbp fragments have been reported earlier³⁷. Our finding of the additional 8.5-kbp fragment could be because we used a molecular probe representing the com-

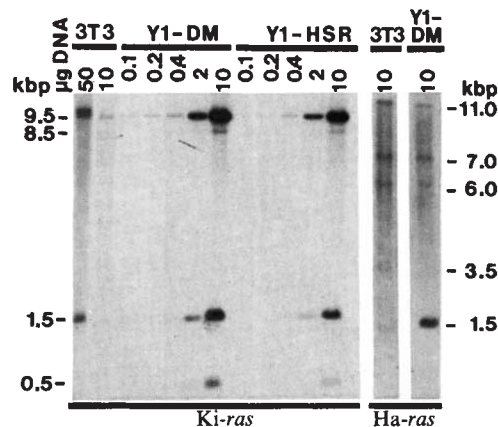


Fig. 4 Analysis of c-Ki-ras in the DNA of Y-DM and Y1-HSR cells. DNA from NIH 3T3 cells served as control. For isolation of high molecular weight DNA, approximately 10^8 cells were lysed with 0.5% SDS in 0.1 M NaCl, 20 mM EDTA, 50 mM Tris-HCl pH 8.1 (total volume ~ 150 ml), and the cellular proteins were hydrolysed with $500 \mu\text{g ml}^{-1}$ proteinase K for 1 h at 37°C . The solution was extracted twice with phenol, and twice with an equal volume of butanol-propanol (7:3). Nucleic acids were precipitated with 3 volumes of ethanol, washed in absolute ethanol and dried in a vacuum. The nucleic acids were redissolved in 1 mM EDTA, 10 mM Tris-HCl, pH 8.1 (TE), and RNA was hydrolysed with $100 \mu\text{g ml}^{-1}$ pancreatic ribonuclease A (Sigma) at 37°C for 1 h. Treatment with proteinase K, extraction with phenol and with butanol-propanol, and ethanol precipitation was performed as above. The DNA was dissolved and stored at 4°C in TE. Aliquots were digested with restriction endonuclease *EcoRI*, fractionated by electrophoresis through a 1% agarose gel and transferred to nitrocellulose paper. The filter was incubated with 5×10^6 c.p.m. of ^{32}P -labelled probe prepared as described for Fig. 1. Conditions for hybridization and washing of filters were as described for Fig. 1, except that 50% formamide was used for c-Ki-ras. kbp, Kilobase pairs. Sizes were calculated using λ phage DNA cleaved with restriction endonuclease *HindIII* as standards.

plete coding region for p21^{c-Ki-ras}, in contrast to the earlier studies that used a probe containing an incomplete coding region^{37,38}. From these and other analyses with restriction endonucleases, we conclude that no major rearrangements have disturbed the topography of c-Ki-ras.

We then determined whether the oncogene *c-Ha-ras*, which is related to *c-Ki-ras*, is also amplified in the Y1 cells. Southern blotting analyses revealed five bands corresponding to fragments of 11.0 kbp, 6.0 bp, 5.0 kbp, 3.5 kbp and 1.5 kbp (Fig. 4). Except for the 1.5-kbp fragment, which probably contains the region of homology shared between *Ki-ras* and *Ha-ras*³⁷, the intensity of the signal obtained for the bands was not different between the NIH 3T3 and the Y1 cells. We therefore conclude that *c-Ki-ras* is amplified, whereas *c-Ha-ras* is not.

Localization of c-Ki-ras in DMs

DMs can be isolated as enriched preparations from Y1-DM cells by a procedure described and validated previously³⁶. To determine whether the DMs in Y1-DM cells carry c-Ki-ras, we analysed DNA extracted from two independent DM preparations. Hydrolysis of the DNAs with *Eco*RI produced fragments of 9.5, 8.5, 1.5 and 0.5 kbp that hybridized with Ki-ras probe (Fig. 5), results identical to those obtained with total DNA from Y1 and NIH 3T3 cells (see above). The amounts of DM DNA required to obtain these results were smaller than those required with total cellular DNA, in rough proportion to our estimates of the extent of amplification of c-Ki-ras in Y1-DM cells. Hybridization with Ha-ras probe detected the 1.5-kbp fragment that shares homology with Ha-ras and Ki-ras, but none of the other c-Ha-ras fragments was apparent (Fig. 5). These results document the enrichment of the preparations of DMs for specific nucleotide sequences and indicate that DMs in Y1-DM cells carry the amplified copies of c-Ki-ras.

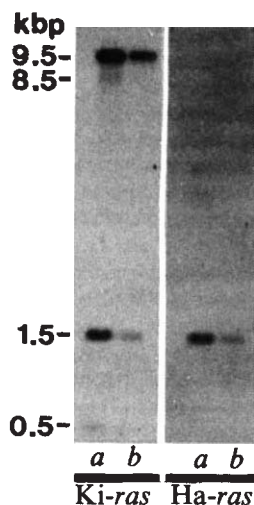


Fig. 5 Analysis of *c-Ki-ras* and *c-Ha-ras* in DNA isolated from DM preparations of Y1-DM cells. Preparation of DMs by differential centrifugation has been described elsewhere³⁶. Compared to unfractionated metaphase chromosome preparations, the preparation is enriched ~100-fold for DM particles, as determined by light microscopic procedures. DNA was extracted from two independent preparations of DMs. Approximately 100 ng DNA of each preparation (*a* and *b*) were digested with restriction endonuclease *EcoRI* and analysed by electrophoresis in agarose gels and by Southern blotting as described for Fig. 4.

Previous studies demonstrated that Y1-DM cells contain an average of 60 DMs per cell³². We have found that *c-Ki-ras* is amplified ~60-fold in Y1-DM cells. It seems likely that each DM carries a single copy of *c-Ki-ras* and may represent a single unit of amplified DNA.

Localization of *c-Ki-ras* to the HSR

The HSR in the Y1-HSR subline is contained in a large marker chromosome, the origin of which is not clear³⁹ (Fig. 6*a*). [A second HSR previously reported to be in another marker chromosome in the Y1-HSR subline³⁹ has been identified in more recent work as a novel euchromatic region produced by complex translocations (D.G., unpublished results).] The distinctive appearance of the HSR marker chromosome (including its size and the presence of a short arm) allows easy identification, even in unbanded chromosome preparations. In order to establish whether the HSR contains *c-Ki-ras*, we hybridized ³H-labelled *v-Ki-ras* probe to spreads of metaphase chromosomes. Autoradiographs revealed a dense cluster of grains along the HSR (Fig. 6*b*). More detailed counts of 15 metaphase spreads gave an average number of 10.2 (range 3–15) grains over the HSR, compared to an average of 0.27 grains over the other chromosomes. We therefore conclude that the HSR contains amplified *c-Ki-ras*.

Conclusions

By deliberately selecting tumour cells containing either DMs or HSRs for analysis, we have uncovered amplification of two cellular oncogenes—*c-myc* in the neuroendocrine tumour cells of the COLO 320 line³¹ and *c-Ki-ras* in the Y1 line of mouse adrenocortical tumour cells as reported here. In both instances, amplification of the cellular oncogene is accompanied by 30- to 60-fold enhancement of expression. Since we encountered amplification of a known oncogene in both of the cell lines we chose to examine first, it seems likely that additional examples of oncogene amplification will emerge as more tumours containing DMs or HSRs are studied.

Gene amplification in eukaryotic cells is thought to result from a spontaneous and random anomaly of DNA replication. For example, it has been estimated that the gene encoding dihydrofolate reductase in the mouse undergoes unsolicited amplification once in every 10⁶ cell divisions⁴⁰. The

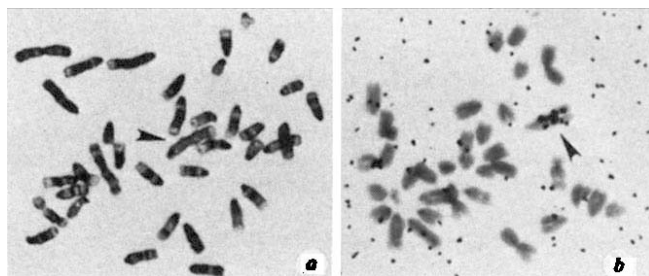


Fig. 6 Chromosomal localization of *c-Ki-ras* by *in situ* hybridization. Exponentially growing cells were exposed to 0.5 $\mu\text{g ml}^{-1}$ colcemide (Gibco) for 60 min, collected by trypsinization, resuspended in 0.075 M KCl for 5 min, fixed in methanol-acetic acid (3:1), spread onto slides and dried. *a*, Trypsin-Giemsa-banded metaphase spread of Y1-HSR cells. The arrow indicates the marker chromosome carrying the HSR. *b*, *In situ* hybridization of ³H-labelled *v-Ki-ras* probe to a metaphase spread prepared from Y1-HSR cells (exposure time 4 days). *In situ* hybridization was performed as follows. Chromosome preparations were treated with pancreatic ribonuclease A (Sigma) at 100 $\mu\text{l min}^{-1}$ in 2 $\times$ SSC at 37°C for 1 h, rinsed well in several changes of 2 $\times$ SSC, and dehydrated through an ethanol series (70%, 80%, 95%). The chromosomal DNA was denatured by incubating the slides at 70°C for 2 min in 70% formamide, 2 $\times$ SSC, pH 7.2. This was followed by quick cooling in 70% ethanol and dehydration through another ethanol series. The *SsrII-EcoRI* fragment of clone HiHi-3 containing the *v-Ki-ras*³⁷ was labelled by nick-translation using [³H]dCTP (51 Ci mmol⁻¹; Amersham), [³H]dATP (34 Ci mmol⁻¹) and [³H]dTTP (53 Ci mmol⁻¹) to a specific activity of 1–2 $\times$ 10⁷ c.p.m. μg^{-1} . Hybridizations were carried out for 15 h at 37°C, using 10⁶–10⁷ c.p.m. per slide in 50–100 μl of 2 $\times$ SSC, 50% formamide, 10% dextran sulphate, 50 mM Tris-HCl, pH 7.6, 0.3 mg ml⁻¹ salmon sperm DNA. The slides were washed in 50% formamide, 2 $\times$ SSC, pH 7 at 39°C for 15 min, 2 $\times$ SSC at room temperature (several changes over 1 h), dehydrated in ethanol and air-dried. They were then exposed to Kodak NTB-2 emulsion for 1–4 days at 4°C, and developed for 2.5 min in Kodak D-19. The chromosomes were stained with Giemsa.

amplification becomes apparent, however, only if the cells in which it occurs are subjected to selective pressure that favours survival of cells bearing the amplified gene. We presume that cellular oncogenes are also amplified spontaneously, but we have no knowledge of the selective pressures or survival advantage that might permit emergence of cells containing an amplified oncogene.

What is responsible for the enhanced expression of *c-Ki-ras* in Y1 tumour cells? Gene amplification offers an obvious explanation because it provides additional DNA template for the synthesis of RNA. In the human leukaemia cells of the HL-60 line, an ostensibly normal allele of *c-myc* is amplified, and expression of the gene is augmented proportionately^{29,30}. Amplification and enhanced expression of *c-myc* have also been found in the neuroendocrine tumour cells of the COLO 320 line³¹. In this instance, however, the amplified *c-myc* takes two forms: an apparently normal configuration, and a configuration that has suffered a major rearrangement of DNA within the 5' end of *c-myc* (M.S. *et al.*, in preparation). It is therefore possible that the rearrangement preceded the amplification of *c-myc* and is principally or entirely responsible for enhanced expression. Recent findings have shown that, in the absence of gene amplification, rearrangements in the vicinity of cellular oncogenes—such as chromosomal translocations^{18,19} or insertions of retroviral DNA^{2,3}—can enhance expression of the genes. The amplified *c-Ki-ras* in Y1 cells appears normal on the basis of initial analyses with restriction endonucleases and molecular cloning (present report and M.S., unpublished results), but more detailed structural analyses and functional tests by gene transfer will be required to decisively test the role of amplification in enhanced expression of the gene.

Has gene amplification contributed to the genesis or maintenance of the Y1 tumour, and if so, has the enhanced expression of *c-Ki-ras* figured in the contribution? The survival of amplified

DNA in the form of DMs implies that the amplification or its consequences (such as enhanced gene expression) has conferred a selective advantage on the cells bearing the DMs. For this reason alone, it seems reasonable to argue that the amplification responsible for the DMs (and HSRs, which are stable in the absence of selective pressure) may have been an essential event in the genesis of the Y1 tumour. The argument does not mandate a role for c-Ki-ras, however, because the domain of amplified DNA in Y1 cells is undoubtedly larger than c-Ki-ras. Genes other than c-Ki-ras might have provided the selective advantage that permitted the survival of the DMs.

Several points suggest that the appearance of c-Ki-ras in the amplified DNA of tumour cells is more than an accident. First, cellular oncogenes recognized as transforming genes of retroviruses have appeared with notable consistency amongst the amplified DNA of tumour cells. Three independent examples are now available (in addition to this report, see refs 29–31). Given the limited number of identified cellular oncogenes (18 at the latest count) and the vast complexity of mammalian DNA, this consistency is provocative. Second, the tumorigenic potential of both rat and human c-Ha-ras (a cellular oncogene

closely related to c-Ki-ras) has been demonstrated experimentally. Molecular clones of c-Ha-ras can transform rodent cells to neoplastic growth if the cloned genes are first joined to a regulatory element from retroviruses that enhances expression of the cellular oncogene^{41,42}. Third, the active transforming gene identified in the DNA of several human tumours and cell lines by gene transfer has proven to be homologous to Ki-ras^{10,11}. These results suggest that the amplification and enhanced expression of cellular oncogenes may contribute to the maintenance of the neoplastic phenotype. The apparent prevalence of cellular oncogenes in amplified DNA of different tumours provides further circumstantial evidence that these genes may be frequent effectors of tumorigenesis.

We thank Ibrahim Tuet for maintenance of cell cultures and J. Marinos for typing the manuscript. This work was supported by grants from the NCI (J.M.B., H.E.V., D.G.) and the American Cancer Society (J.M.B., H.E.V.). M.S. is a Heisenberg-fellow of Deutsche Forschungsgemeinschaft, K.A. is a fellow of the NIH Fogarty International Center. We thank M. Yoshida, H. Temin, I. Verma, S. Aaronson, E. Scolnick, C. Sherr and C. Hammond for molecular clones of oncogenes.

Received 28 February; accepted 21 April 1983.

1. Bishop, J. M. A. *Rev. Biochem.* (in the press).
2. Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475–480 (1981).
3. Payne, G. S., Bishop, J. M. & Varmus, H. E. *Nature* **295**, 209–217 (1982).
4. Noori-Daloii, M. R., Swift, R. A., Kung, H., Crittenden, L. B. & Witter, R. L. *Nature* **294**, 574–576 (1981).
5. Nusse, R. & Varmus, H. E. *Cell* **31**, 99–109 (1982).
6. Varmus, H. E. *Science* **216**, 812–820 (1982).
7. Weinberg, R. A. *Adv. Cancer Res.* **36**, 149–163 (1982).
8. Cooper, G. M. *Science* **218**, 801–817 (1982).
9. Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474–478 (1982).
10. Pulciani, S. et al. *Nature* **300**, 539–542 (1982).
11. Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637–3640 (1982).
12. Tabin, C. J. et al. *Nature* **300**, 143–149 (1982).
13. Reddy, E. P., Reynolds, R. N., Santos, E. & Barbacid, E. *Nature* **300**, 149–152 (1982).
14. Taparowsky, E. et al. *Nature* **300**, 762–765 (1982).
15. Taub, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7837–7841 (1982).
16. Crews, S., Barth, R., Hood, L., Prehn, J. & Calame, K. *Science* **218**, 1319–1321 (1982).
17. Sheing-Ong, G. L. C., Keath, E. J., Piccoli, S. P. & Cole, M. D. *Cell* **31**, 443–452 (1982).
18. Marcu, K. B. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 519–523 (1983).
19. Erikson, J., Ar-Rushdi, A., Drwings, H., Nowell, P. & Croce, C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 820–824 (1983).
20. Adams, J., Gerondakis, S., Webb, E., Corcoran, L. M. & Cory, S. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

21. Dalla-Favera, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7824–7827 (1982).
22. de Klein, A. et al. *Nature* **300**, 765–767 (1983).
23. Rechavi, G., Givol, D. & Canaani, E. *Nature* **300**, 607–611 (1982).
24. Schimke, R. T., Kaufman, R. J., Alt, F. W. & Kellems, R. F. *Science* **202**, 1051–1055 (1978).
25. Barker, P. E. *Cancer Genet. Cytogenet.* **5**, 81–94 (1982).
26. Cowell, J. K. A. *Rev. Genet.* **16**, 21–59 (1982).
27. Nunberg, J. N., Kaufman, R. J., Schimke, R. T., Urlaub, G. & Chasin, L. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5553–5556 (1978).
28. Kaufman, R. J., Brown, P. C. & Schimke, R. T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5669–5673 (1979).
29. Collins, S. & Groudine, M. *Nature* **298**, 679–681 (1982).
30. Dalla-Favera, R., Wong-Staal, F. & Gallo, R. *Nature* **299**, 61–63 (1982).
31. Alitalo, K., Schwab, M., Lin, C. C., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1707–1711 (1983).
32. George, D. L. & Francke, U. *Cytogenet. Cell Genet.* **28**, 217–226 (1980).
33. Ellis, R. W., DeFeo, D., Furth, M. E. & Scolnick, E. M. *Molec. cell. Biol.* **2**, 1339–1345 (1982).
34. Furth, M. E., Davis, L. J., Fleurdelys, B. & Scolnick, E. M. *J. Virol.* **43**, 294–304 (1982).
35. Wittingham, M. C., Pastan, I., Shih, T. Y. & Scolnick, E. M. *Cell* **19**, 1005–1014 (1980).
36. George, D. L. & Powers, V. E. *Cell* **24**, 117–123 (1981).
37. Ellis, R. W. et al. *Nature* **292**, 506–511 (1981).
38. Tsuchida, N., Ryder, T. & Ohtsubo, E. *Science* **217**, 937–938 (1982).
39. George, D. L. & Powers, V. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1597–1601 (1982).
40. Varshavsky, A. *Cell* **25**, 561–572 (1981).
41. DeFeo, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3328–3332 (1981).
42. Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. *Nature* **292**, 479–483 (1982).

LETTERS TO NATURE

SS433 modelled as an aligned magnetic neutron star

C. Shukre* & R. N. Manchester

Division of Radiophysics, CSIRO, PO Box 76, Epping, New South Wales 2121, Australia

D. A. Allen

Anglo-Australian Observatory, PO Box 296, Epping, New South Wales 2121, Australia

SS433 is a variable X-ray, optical and radio source lying centrally within the supernova remnant W50 (see review in ref. 1). The notoriety of SS433 rests on the unique, opposing pair of collimated gaseous jets of opening angle $\sim 5^\circ$ and velocity $0.27c$ (ref. 2). It is a binary system of period 13 days (ref. 3). The jets describe a cone of half-angle 20° with period 164 days

(refs 4, 5) which is usually interpreted as a precession. We propose here that SS433 is a neutron star whose rotational and magnetic axes are almost aligned. Alignment provides collimation of the jets which originate in supercritical accretion bursts. Precession of the neutron star is invoked to account for the observed 164-day periodicity. The corotating magnetosphere gives the star a quadrupole moment so that classical precession occurs in the gravitational field of the accretion disk.

In any model of SS433 the collimation and precession of the jets are the principal features to be explained. Interpretations generally invoke accretion from an ordinary star on to a compact object. The accretion is probably by Roche lobe overflow⁶, and it is likely that an accretion disk exists. Material in the jets is emitted in a series of discrete 'bullets'⁷ which are probably related to outbursts in the radio and X-ray emission and the existence of discrete components in very large array (VLA)⁸ and very long baseline interferometry (VLBI) (ref. 9; also Romney, Schilizzi, Spencer and Fejes, in preparation) maps. The high ejection velocity is symptomatic of an object with a surface escape velocity approaching the speed of light.

The accreting object has often been modelled as a neutron star, but the fact that neutron stars normally have magnetic fields $\sim 10^{12}$ G has generally been ignored. We explore here

* Permanent address: Raman Research Institute, Bangalore, India.

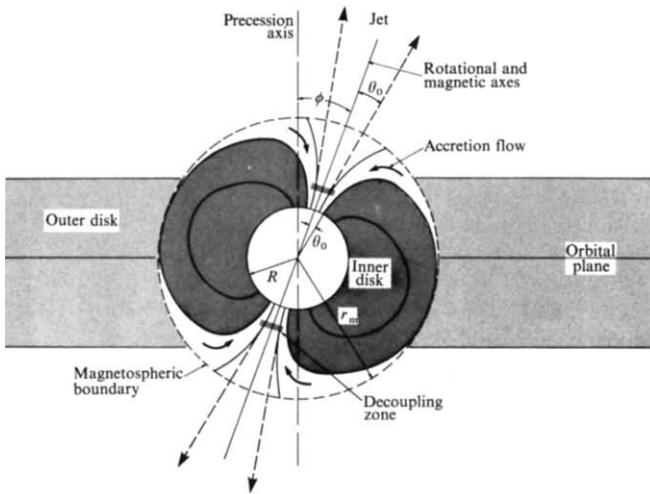


Fig. 1 Schematic diagram of the aligned magnetic neutron star model for SS433. For clarity, radii have not been drawn to scale, the angle θ_0 (the half width of the polar cap and of the jet) has been increased and the angle χ between the rotation and magnetic axes has been taken as zero.

the case of accretion on to a magnetic neutron star, and are able to account for many of the observed properties of SS433. A list of symbols used subsequently, together with their assumed or derived values, is given in Table 1.

The accretion flow is greatly modified by the stellar magnetic field. Within the magnetosphere, which is bounded by the Alfvén surface where the magnetic and plasma energy densities are equal, matter preferentially moves along field lines and hence is channelled towards the magnetic poles. Similarly the magnetic field in the polar regions provides collimation of the ejected gas (see refs 10, 11). The Alfvén surface is roughly spherical and of radius r_m , typically 3×10^8 cm. If the magnetosphere accretes isotropically, most of the material will fall in annular regions around each polar cap, as illustrated in Fig. 1. The polar cap, defined by the open field lines, has a halfwidth of θ_0 . Typically $\theta_0 \sim 3^\circ$, and almost 90% of the accreted material lands in the annulus between $\theta_0/2$ and θ_0 . This effect is enhanced if the accretion flow is from a disk of thickness $< 2r_m$. Consequently, a natural nozzle is formed by the low-density zone around the magnetic axis, allowing supercritical accretion to drive the outflow. We take the halfwidth of this nozzle at the stellar surface to be $\theta_0/2$.

The accretion rate is too high for thermonuclear flashes to occur¹², and in any case the dominant energy supply is gravitational. The burst-like activity of SS433 must be attributed to instabilities in the accretion flow. The recurrence time for instabilities, t_R , as inferred from the separation between radio blobs seen in VLBI⁹, is about 2 days. A possible instability mechanism is the 'magnetospheric gate' of Lamb *et al.*¹³, although it is not clear that a recurrence time scale as long as 2 days can be accommodated without some modification of the model.

Once super-Eddington luminosity is generated matter will be driven out through the nozzle. A similar idea has been discussed¹⁴ in the context of Her X-1. Close to the star the kinetic energy density of the ejected matter must be lower than the magnetic energy density so that collimation results. Since the magnetic energy density falls as the inverse sixth power of radial distance whereas the kinetic energy density of the outflowing gas decreases slower than the inverse fourth power of radius, the gas will decouple from the magnetic field within a few stellar radii. The angular halfwidth of the emerging jet is defined at the decoupling radius and is not very sensitive to the initial kinetic energy. Typically it is close to θ_0 in agreement with observations.

In order that successive ejections define an observable jet it is crucial that the magnetic and rotational axes be nearly aligned. This ensures that the collimation axis remains relatively stable for times long compared with the rotation period of the star. If the two axes were misaligned by an angle χ , then for $\chi > \theta_0$ the solid angle into which material was ejected would be increased by a factor $4/(\theta_0 \sin \chi) \sim 100$, making the jets much more difficult to detect. They would probably be absorbed into the accretion disk.

The requirement of alignment also provides a natural explanation for the present unique status of SS433. Pulsar birthrates are estimated to be comparable with or even greater than supernova rates¹⁵, indicating that most neutron stars are born with the magnetic and rotation axes misaligned. Assuming random orientation at birth, the probability of $\chi < \theta_0 \sim 3^\circ$ is only a few per cent. Since there are ≈ 100 known or suspected X-ray binary systems in the Galaxy¹⁶, one would expect only a few systems like SS433. Other factors such as short lifetime or the requirement of rapid mass transfer may further limit the number of observed systems.

Consider a mass ΔM of infalling material. The gravitational energy released exceeds the Eddington luminosity so that a fraction $x\Delta M$ is ejected by radiation pressure. For a neutron star of mass M and radius R the energy released is

$$\Delta E = (GM/R)(1-x)\Delta M = \frac{1}{2}(1-x)\Delta M V_{es}^2 \quad (1)$$

where $V_{es} = (2GM/R)^{1/2}$ is the escape velocity and we have assumed unit conversion efficiency of gravitational potential energy.

We equate this to the energy supplied by radiation which ejects a portion of the infalling gas

$$\Delta E = cx\Delta M(V_{ej} + V_f) \quad (2)$$

where V_{ej} and V_f are the magnitudes of the ejection and infall velocities near the neutron star's surface. Hence

$$(1-x)V_{es}^2 = 2xc(V_{ej} + V_f) \quad (3)$$

Far from the star the ejected matter moves with velocity V_j and

$$V_j^2 = V_{ej}^2 - V_{es}^2 \quad (4)$$

Thus

$$(1-x)/x = 2c[(V_j^2 + V_{es}^2)^{1/2} + V_f]/V_{es}^2 \quad (5)$$

Now, $V_j = 0.27c$ and $V_{es} \sim 0.5c$ for $M \sim 1M_\odot$, $R \sim 10^6$ cm. Taking $V_f \sim V_{es}$ we find $x \sim 0.1$.

If $\dot{M}_j$ is the mass-loss rate in the jets, we have $\dot{M}_j = [x/(1-x)]\dot{M}$, where $\dot{M}$ is the average accretion rate on to the neutron star. Estimates of $\dot{M}$ (ref. 17) are 10^{-4} – $10^{-5} M_\odot \text{ yr}^{-1}$, and of $\dot{M}_j$ (ref. 18) 10^{-6} – $10^{-7} M_\odot \text{ yr}^{-1}$. We adopt $\dot{M} = 10^{-5} M_\odot \text{ yr}^{-1}$ and $\dot{M}_j = 10^{-6} M_\odot \text{ yr}^{-1}$. Even for isotropic accretion on to a neutron star, this rate is far in excess of that required to exceed the Eddington luminosity; since in our case most of the incident material lands on $< 10^{-2}$ of the neutron star's surface, the local accretion rate is highly supercritical. For a burst recurrence time $t_R \sim 2$ days, $\Delta M \sim 5 \times 10^{-8} M_\odot \sim 10^{26}$ g.

We may rearrange equation (5) to give

$$V_j = (c/8x)(1-9x)^{1/2}(1-x)^{1/2} \quad (6)$$

where we have again used $V_f = V_{es} = 0.5c$. The observed constancy of V_j then indicates that x is constant to better than 1%. This is a constraint on the accretion instability mechanism.

We require a precession mechanism for the neutron star. Free precession would not suffice since it would destroy the alignment of the magnetic and rotation axes. Precession forced by the companion, both in relativistic and newtonian cases, is ineffective unless the binary period is of the order of minutes^{10,19,20}. However, the presence of a strong magnetic field suggests a precession mechanism. A trickle of gas is expected to percolate in through the Alfvén surface (see ref. 13).

There it is supported by magnetic and rotational forces, and may accumulate to a substantial mass. This material is coupled to the neutron star by the magnetic field and co-rotates with it. Thus a quadrupole component is generated in the moment of inertia of the system star-plus-magnetosphere. This system can now be forced to precess by a mass in orbit around it. The companion is still too distant but the outer part of the accretion disk could be effective.

The concept of an inner disk coupled to the neutron star by Lense–Thirring forces has already been explored by Sarazin *et al.*²¹. They further proposed that the inner disk-star system experiences Lense–Thirring precession in the field of the outer parts of the accretion disk. This requires the outer disk to be extremely massive and to have large angular momentum. The viscosity parameter, α (ref. 22), of the material must be $\sim 10^{-5}$. In our model the magnetosphere has the role of the inner disk and has a radius ($r_m \sim 3 \times 10^8$ cm) comparable with that of Sarazin *et al.* However, we believe that the newtonian quadrupole precession by the outer disk is a more likely mechanism than Lense–Thirring precession. It is possible that both exist, and since they have opposite signs the larger will dominate. The observational evidence, however, seems not to favour the Lense–Thirring mechanism²³. We analyse the newtonian case.

For a keplerian disk the precession frequency Ω is given by

$$\Omega = (3\Delta I/4\pi I_p)P \cos \phi GM_{od}/a^3 \quad (7)$$

where $\Delta I = I_p - I_e$ and is the difference in the moments of inertia about polar and equatorial axes, P is the neutron star's rotation period, ϕ is the observed half-angle of the precession cone, and M_{od} is an effective mass for the outer disk at a dynamically

equivalent radius a . We have chosen to approximate the integral over the outer disk in this way since we do not know the structure of the disk when an accretion instability controls its inner boundary.

In equation (7), ΔI represents the contribution of the inner disk and can be approximated by $0.25 M_{id} r_m^2$, where M_{id} is the mass of the inner disk. Hence we derive

$$\frac{\Delta I}{I_p} \approx 5 \times 10^{-4} \left[\frac{M_{id}}{M_\odot} \right] \left[\frac{r_m}{3 \times 10^8 \text{ cm}} \right]^2 \left[\frac{10^{45} \text{ g cm}^2}{I_p} \right] \quad (8)$$

and thus

$$\frac{\Omega}{\Omega_{\text{obs}}} \approx 1.5 \times 10^{11} \left[\frac{M_{id}}{M_\odot} \right] \left[\frac{M_{od}}{M_\odot} \right] \left[\frac{10^{45}}{I_p} \right] \left[\frac{r_m}{3 \times 10^8} \right]^{-1} \times \left[\frac{r_m}{a} \right]^3 \left[\frac{P}{s} \right] \quad (9)$$

where Ω_{obs} is the observed precession frequency.

What are reasonable masses for the outer and inner disks? Studies of dwarf novae and symbiotic stars²⁴ suggest that the viscosity parameter $\alpha \sim 1$ in these systems, and rapid energy dissipation results in material passing through the disk in about an orbital period. In SS433, for $M = 10^{-5} M_\odot \text{ yr}^{-1}$ and a 13-day period, $M_{od} \sim 4 \times 10^{-7} M_\odot$. The mass supported in the inner disk is more uncertain, and we can only take a guess at $M_{id} \sim 10^{-6} M_\odot$. For these parameters we then require that $a/r_m \sim 0.4 P^{1/3}$, where P is in seconds.

Now, a/r_m must considerably exceed unity, and this condition is most easily met if $P \geq 10^2$ s. Failing this, we must postulate larger masses for the inner and outer disks. The latter will follow if $a < 1$. However, α need not be as low as 10^{-5} , the value required for the Lense–Thirring mechanism. We therefore suggest that either $\alpha \sim 1$ in the accretion disks and that the neutron star rotation period is relatively long, or $\alpha \ll 1$, in which case the rotation period could be shorter.

Analysis of the jitter in the observed jet radial velocities^{6,25} shows that it is direction rather than ejection speed of the jet which varies. It would seem difficult to accommodate this in any model where the jet direction is defined by precession of a neutron star. However, in the present model the nozzle is defined by the accretion flow to the polar annulus, and this will not necessarily possess perfect circular symmetry. Moreover, if the burst duration is shorter than the rotation period of the neutron star, then a small misalignment χ will generate an apparently random variation in the jet direction.

We thank Richard Schilizzi for spurring our interest in SS433 and for useful discussions.

Received 31 January; accepted 8 April 1983.

- Beer, P. *Vistas Astr.* **25**, Pts 1, 2 (1981).
- Margon, B., Grandi, S. A. & Downes, R. A. *Astrophys. J.* **241**, 306–315 (1980).
- Crampton, D., Cowley, A. P. & Hutchings, J. B. *Astrophys. J. Lett.* **235**, L131–L135 (1980).
- Abell, G. O. & Margon, B. *Nature* **279**, 701–703 (1979).
- Hjellming, R. M. & Johnston, K. J. *Nature* **290**, 100–107 (1981).
- van den Heuvel, E. P. J. *Vistas Astr.* **25**, 95–108 (1981).
- Grandi, S. *Publ. astr. Soc. Pacif.* **94**, 80–86 (1982).
- Gilmore, W. S., Seaquist, E. R., Stocke, J. T. & Crane, P. C. *Astr. J.* **86**, 864–870 (1981).
- Schilizzi, R. T. *et al. IAU Symp.* No. 97, 205–206 (1982).
- Shaham, J. *Astrophys. Lett.* **20**, 115–117 (1980).
- Fabian, A. G. *Mon. Not. R. astr. Soc.* **192**, 11P–14P (1980).
- Ayasli, S. & Joss, P. C. *Astrophys. J.* **256**, 637–665 (1982).
- Lamb, F. K., Fabian, A. C., Pringle, J. E. & Lamb, D. Q. *Astrophys. J.* **217**, 197–212 (1977).
- Trümper, J., Morfill, G., Bodenheimer, P. & Tenorio-Tagle, G. *Rep. Workshop on Accreting Neutron Stars*, Garching (1982).
- Taylor, J. H. & Manchester, R. N. *Astrophys. J.* **215**, 885–896 (1977).
- Holt, S. S. & McCray, R. A. *Rev. Astr. Astrophys.* **20**, 323–365 (1982).
- van den Heuvel, E. P. J., Ostriker, J. P. & Peterson, J. A. *Astr. Astrophys.* **81**, L7–L10 (1980).
- Hjellming, R. M. & Johnston, K. J. *IAU Symp.* No. 97, 197–204 (1982).
- Martin, P. G. & Rees, M. *Mon. Not. R. astr. Soc.* **189**, 19P–22P (1979).
- Begelman, M. C., Sarazin, C. L., Hatchett, S. P., McKee, C. F. & Arons, J. *Astrophys. J.* **238**, 722–730 (1980).
- Sarazin, C. L., Begelman, M. C. & Hatchett, S. P. *Astrophys. J. Lett.* **238**, L129–L132 (1980).
- Pringle, J. E. A. *Rev. Astr. Astrophys.* **19**, 137–162 (1981).
- Katz, J. I., Anderson, S. F., Margon, B. & Grandi, S. A. *Astrophys. J.* **260**, 780–793 (1982).
- Bath, G. T. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **201**, 345–355 (1982).
- Margon, B. *Ann. N.Y. Acad. Sci.* **375**, 403–413 (1981).

Table 1 Symbols and parameters

		Adopted value	Derived value
a	Characteristic radius of outer disk	$10^{9 \pm 0.5} \text{ cm}$	
I_p	Polar moment of inertia of neutron star	10^{45} g cm^2	
M_{id}	Mass of inner accretion disk	$10^{-6} M_\odot$	
M_{od}	Mass of outer accretion disk		$\geq 4 \times 10^{-7} M_\odot$
M	Mass of neutron star	$1 M_\odot$	
$\dot{M}$	Mean accretion rate to neutron star	$10^{-5} M_\odot \text{ yr}^{-1}$	
$\dot{M}_j$	Mean mass loss rate in jets	$10^{-6} M_\odot \text{ yr}^{-1}$	
ΔM	Mass of accreted gas causing a burst		$3 \times 10^{-8} M_\odot$
P	Neutron star rotation period		$10^{1 \pm 1} \text{ s}$
r_m	Radius of magnetosphere	$3 \times 10^8 \text{ cm}$	
R	Radius of neutron star	10^6 cm	
t_R	Recurrence time of burst	2 days	
V_{ej}	Ejection velocity at stellar surface		$0.57c$
V_{es}	Escape velocity from neutron star	$0.5c$	
V_f	Free-fall velocity from magnetosphere	$= V_{es}$	
V_j	Observed jet velocity	$0.27c$	
α	Disk viscosity parameter		$\ll 1$
x	Fraction of ΔM which is ejected		0.1
θ_0	Half opening angle at polar cap	$\sim 3^\circ$	
ϕ	Half-angle of precession cone	20°	
χ	Angle between magnetic and rotation axes		$\leq \theta_0$
Ω	Angular precession frequency	(164 days^{-1})	

Very long baseline interferometry at a wavelength of 3.4 mm

A. C. S. Readhead*, C. R. Masson*, A. T. Moffet*,
T. J. Pearson*, G. A. Seielstad*, D. P. Woody*,
D. C. Backer†, R. L. Plambeck†, W. J. Welch†,
M. C. H. Wright†, A. E. E. Rogers‡, J. C. Webber‡,
I. I. Shapiro§, J. M. Moran||, P. F. Goldsmith¶,
C. R. Predmore¶, L. Bååth* & B. Rönnäng*

* Owens Valley Radio Observatory, California Institute of Technology, Pasadena, California 91125, USA

† Radio Astronomy Laboratory, University of California, Berkeley, California 94720, USA

‡ Haystack Observatory, Westford, Massachusetts 01886, USA

§ Department of Earth and Planetary Science, Massachusetts Institute of Technology, Massachusetts 02139, USA

|| Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA

¶ Physics and Astronomy Department, University of Massachusetts, Amherst, Massachusetts 01003, USA

* Onsala Space Observatory, S-430 34 Onsala, Sweden

In recent years the angular distribution of intensity of the compact cores of quasars and radio galaxies has been mapped in detail by very long baseline interferometry (VLBI) arrays^{1,2}. These observations have shown that asymmetric radio structure is common in active nuclei³, and that a high proportion of the brightest sources exhibit superluminal motion^{4,5}. The sources generally have an elongated structure, or jet, which dominates at long centimetre wavelengths, and a compact core at one end of the jet which dominates at short centimetre wavelengths. The association of the core components with the shortest wavelengths, and the most compact features, leads us to link core radiation with the unidentified central engines in these energetic systems. We report here the first detection of interference fringes by VLBI at a wavelength of 3.4 mm (89 GHz). The success of these observations demonstrates the feasibility of millimetre VLBI and opens this important wavelength range to mapping with a resolution of 10^{-4} arc s. The compact radio source in the nucleus of 3C84 (NGC1275) was observed for a 9-h period on the 485-km ($1.4 \times 10^8 \lambda$) baseline between the Owens Valley Radio Observatory and the Hat Creek Observatory. These observations show that the structure is more compact at 89 GHz than at 22 GHz.

There are some compelling astrophysical reasons for attempting VLBI at the highest possible frequencies: (1) The resolution provided by millimetre VLBI enables us to study nearby active galactic nuclei on a scale $\sim 10^{17}$ cm, which is comparable with the predicted size of an accretion disk around a $10^9 M_{\odot}$ black hole⁶; it also enables us to study the cores of many quasars on a scale < 1 pc. (2) High-frequency observations can be used to probe the optically thick cores and reveal structures associated with the central engine. (3) The radio emission regions at frequencies > 20 GHz are probably comparable in size to the broad emission line regions. Thus a combination of radio structure information with optical line and continuum spectral information could help to determine the nature of the energy source.

Over the past few years there have been several technological advances which now make it possible to do continuum VLBI at millimetre wavelengths. Chief among these are the construction of radio telescopes which operate at millimetre wavelengths, the development of low-noise millimetre receivers, and the development of the broad-band Mark III VLBI recording system and processor⁷. In view of the small sizes (< 20 m) of most millimetre wavelength radio telescopes the widest possible recording bandwidths must be used. This need places heavy demands on timing and fringe rate accuracy.

A successful experiment at 89 GHz was conducted on 21–23 October 1981 between the Owens Valley Radio Observatory

in Big Pine, California (OVRO) and the Hat Creek Radio Observatory in Cassel, California (HCRK). The OVRO–HCRK baseline is 485 km, or 1.4×10^8 wavelengths. Since both observatories also have centimetre wavelength antennas, simultaneous observations at 5 GHz were performed to provide a continuous check of the relative epoch and rate errors of the clocks. The 5 GHz observations also provide a useful limit on the uncertainties due to baseline errors since the offsets at each site between the millimetre and centimetre antennas are known to within a few centimetres. The 5-GHz observations were recorded on two tracks of the Mark III recorder, leaving 26 tracks (52 MHz total bandwidth) for the 89-GHz observations. The 5-GHz observations and the real-time correlation system at the Smithsonian Astrophysical Observatory were used to determine the relative clock epoch and rate offsets. The observations were correlated on the Mark III VLBI Processor at the Haystack Observatory.

Fringes were detected at 89 GHz on all observations of 3C84 (NGC1275; 0316+413), with the consistency of fringe rates leaving no doubt about the reality of the detections. This object was observed for more than 9 h (16 scans of 13 min duration each) on 21 October and for three scans, each also 13 min in duration, on 22 October, and one 6-min scan on 23 October. The sensitivity of the observations was limited by a short coherence time of only a few seconds. For the other objects observed (3C273 and 3C345), no fringes were detected.

The fringe amplitude was measured as a function of coherent integration period and was found to decrease for all periods > 0.5 s. The decrease was 40% for an increase in integration interval from 0.5 to 10 s; this latter time was used in the reduction of all the data as it yielded the highest overall signal-to-noise ratio. Although there are several factors that limit coherence time, the most important was probably the phase noise introduced by a 5-MHz crystal oscillator in the maser-crystal oscillator combination at Hat Creek. Tropospheric phase fluctuations were negligible on the 10-s time scale since the observations were at night, the temperature was below freezing, there was no wind and the sky was completely clear. This conclusion was confirmed by a subsequent 89-GHz VLBI run in May 1982, in which coherence times of several hundred seconds were obtained. Data from this later experiment are still being processed. In VLBI observations⁸ of SiO maser sources at 43 GHz the fringe visibility amplitudes decreased by a factor of 0.7 in 300 s.

Correlated flux densities were derived in the following manner. For each tape pair the correlation coefficients from the 10 s coherent averages were incoherently averaged. These coefficients were corrected for the measured decorrelation between 0.5 and 10 s, as described above, for clipping, for approximations used in the correlation process, and for the bias due to the low signal-to-noise ratio according to the expression⁹:

$$\rho_{\text{corr}} = [\sum (A_i^2 - 2\sigma^2)/n]^{1/2}$$

Where the A_i ($i = 1 \rightarrow n$) are the corrected 10-s correlation coefficients, σ is the r.m.s. value of the noise in the real and imaginary parts of the signal off source, and n is the number of samples. Correlated flux densities, S_{corr} , were then calculated from the values of ρ_{corr} using the ratios of system temperature, T_s , to antenna temperature, T_a , at the two stations:

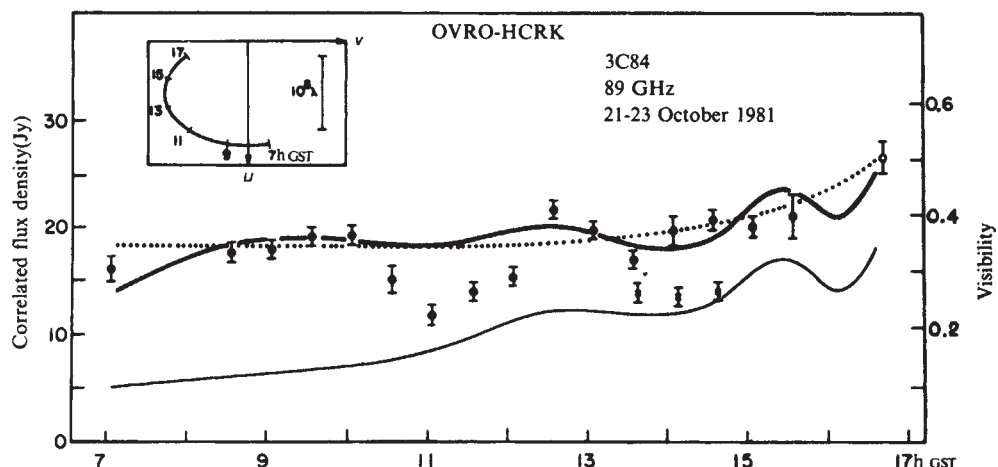
$$S_{\text{corr}} = [T_{sO} T_{sH} T_{aO} T_{aH}]^{1/2} S \rho_{\text{corr}}$$

where subscripts O and H refer to Owens Valley and Hat Creek, respectively, and S is the total flux density.

At Hat Creek the Whitaker 6-m antenna was used on days 294 and 295, and the Philco 6-m antenna was used on day 296. The ratio T_{aH}/T_{aO} was measured at intervals of about 2 h, with an r.m.s. accuracy of $\sim 5\%$, during the observations. The system temperature, T_{sH} , was ~ 400 K throughout the observing period.

At Owens Valley the number two 10.4-m antenna was used for all of the observations. The receiver temperature was measured a few times each day and found to be constant to

Fig. 1 The correlated flux densities observed at 89 GHz on three consecutive days: day 294, ●; day 295, ×; day 296, ○. The (1σ) uncertainties due to random errors are shown by the vertical bars. The 3σ uncertainty in the flux density scale is $\sim 40\%$ (see text). The fits of the three models from Table 1 are also shown: model 1, dotted line; model 2, thin solid line; model 3, thick solid line. The inset shows the variation of the projected baseline length and orientation during the observing period.



within $\sim 20\%$ (3σ) between retuning. The system temperatures were calculated from these values using a sky opacity of 0.08. The total system temperature at Owens Valley was 600 ± 150 K (3σ), assuming equal power in the upper and lower sidebands of the receiver. The aperture efficiency of the number two 10.4-m antenna at OVRO, determined from observations of Mars, is $40\% \pm 10\%$ (3σ). The antenna temperatures at OVRO were estimated using this value and the value of the sky opacity quoted above.

The overall 3σ uncertainty in the Owens Valley and Hat Creek calibrations is $\sim 50\%$ and 20% , respectively. Thus the flux density scale in the VLBI observations could be in error by $\sim 40\%$. In addition to these corrections which were measured, or could be estimated, there may be a loss of correlation from high frequency phase noise in the two local oscillator chains. However, independent oscillator tests, in which the two antennas at each site were locked to independent synthesizers, and the use of these antennas in local interferometer systems suggest that this unknown factor is small.

The total flux density of 3C84 was 53 ± 5 Jy, based on OVRO observations in the week following the VLBI observations. The correlated flux densities (see Fig. 1) show a nearly constant visibility of 0.4 during the 9-h observation on 21 October 1981, when the projected baseline rotated through 87° and changed in length by $45 \times 10^6 \lambda$ or 30% (see insert in Fig. 1). The day-to-day variations are known to arise from tracking errors at OVRO. These variations originated in a software error in the pointing program, which was discovered after these observations. The variations on a time scale of hours are probably due to tracking errors at OVRO and/or small variations due to source structure.

An earlier attempt at 89 GHz VLBI was carried out on 19–21 March 1981 between telescopes at the Owens Valley Radio Observatory, the Five Colleges Radio Astronomy Observatory at Quabbin, Massachusetts, and the Onsala Space Observatory, Sweden. No strong interference fringes were detected in these observations above estimated visibilities of ≤ 0.15 for 3C84 and ≤ 0.35 for 3C345. The sources may be heavily resolved on these long baselines. On the other hand, there may have been technical difficulties at two of the three sites. These difficulties include pointing problems, LO phase noise, and clock epoch and rate errors. A wide search over 10^7 points in fringe rate and delay on each of the three baselines resulted in three apparently significant fringe detections: an 8.4σ detection on 3C84 between Quabbin and OVRO corresponding to a visibility of 0.13; and a simultaneous detection on 3C345 between Quabbin and Onsala (6.4σ) and between Onsala and OVRO (6σ), corresponding to visibilities of 0.3 and 0.35, respectively. If we assume that the noise at each telescope has a gaussian distribution, and that there is no bias in the correlator, we may calculate the probability of observing a random noise peak, in 10^7 trials, at a signal-to-noise ratio greater than or equal to the levels detected here. The probabilities of such a mis-

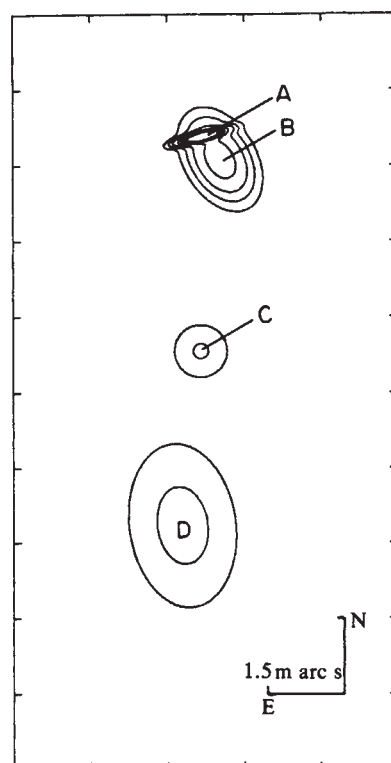


Fig. 2 Four component model of the nuclear radio emitting regions in 3C84 derived from observations at 22 GHz (ref. 10). The linear scale is $0.5 \text{ pc m arc s}^{-1}$, assuming $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The present 89-GHz observations suggest that the compact emission region at this frequency is located in or near component A (see text).

identification of signal are 10^{-6} , 0.02 and 0.1, respectively. Nevertheless the reality of these fringes is questionable since no fringes were detected on adjacent scans. It is conceivable that the high signal-to-noise ratios are caused by some unknown systematic bias in the correlator. It would be straightforward to check this by correlating tapes recorded at different times; however, this would require considerable correlator time, and it has not yet been possible to perform this test owing to the heavy use of the Mark III processor.

It is not possible to make a detailed determination of the 89-GHz structure of 3C84 based on these observations on a single baseline. We have used weighted least squares to estimate the parameters for a simple core-halo model. For this purpose we have used the upper envelope of the data shown in Fig. 1, and ignored the low points in the interval GST 10 h $\rightarrow$ 12 h, since the low points are probably affected by mispointing. We have not included the systematic errors in this analysis, since

Table 1 The structure of 3C84 at 89 GHz—parameters of the models

Model	Component	α_{10}^{22} ($S \propto \nu^\alpha$)	Size (FWHM) (m arc s)		Position angle	Flux density at 89 GHz (Jy)	Mean correlated flux density on OVRO-HCRK baseline (Jy)
			Major axis	Minor axis			
1	Core	—	≤ 0.5		—	28	19
	Halo	—	—		—	25	—
2	A	+1.5	0.65×0.13		105°	6	4
	B	+0.9	1.15×0.7		30°	20	7
	C	—	0.85×0.85		0°	3	0.9
	D	0.0	2.5×1.6		8°	22	0.3
3	A	+1.5	1.0×0.15		80°	19	14
	B	+0.9	1.15×0.7		30°	23	8
	C	—	0.85×0.85		0°	2	0.6
	D	0.0	2.5×1.6		8°	9	0.1

the internal calibration is good enough to enable us to rule out some interesting classes of models. The resultant core-halo model, number 1 in Table 1, gives the dotted line in Fig. 1. The last column of Table 1 shows the correlated flux density of each component alone, averaged over the epoch spanned by the observations. In this model there is a 28 Jy core component ≤ 0.5 m arc s in diameter, and a 25 Jy completely resolved halo. This model gives a χ^2 per degree of freedom of 1.6. A fair fit, which gives a χ^2 per degree of freedom of 2.2, is given by a 0.7 m arc s diameter core of 39 Jy, but core components of larger angular size do not fit the data very well.

It is interesting to compare the present observations with the observed structure at 22 GHz (ref. 10) shown in Fig. 2 and labelled Model 2 in Table 1. This model is derived from 22 GHz VLBI observations made in April 1981, at which time the total 22 GHz flux density was 52 ± 2 Jy. The thin solid curve in Fig. 1 shows the expected visibilities on the OVRO-HCRK baseline at 89 GHz under the assumption that the 89 GHz structure is identical to that in Fig. 2. It is clear from Fig. 1 that the observed visibilities at 89 GHz are considerably larger than the visibilities given by the 22 GHz model. This is consistent with a trend found in centimetre wavelength VLBI: the size of the core component decreases with increasing frequency.

In view of the paucity of data and the uncertainty in the pointing of the 10-m telescope, more detailed model fitting is not justified by the present observations. Nevertheless, note that plausible models, based on model 2 (Fig. 2) but with more flux density in the northern components (A and B) can be found. An example of such a model, with $\chi^2 = 2.2$ per degree of freedom, is given in Table 1 (model 3). The fit is shown by the heavy solid line in Fig. 1. Furthermore, plausible models based on model 2 but with the extra 89 GHz correlated flux density in the southern components (C and D) cannot be made to fit the data. Between 3 and 5 cycles with peak to peak modulations of ≥ 19 Jy in the correlated flux density would be seen unless the flux density of components A+B drops to less than half the value at 22 GHz—which is unlikely since they both have strongly inverted spectra between 10 and 22 GHz. Thus the present observations provide some evidence that the most compact features at 89 GHz are situated in or near components A and B.

The results presented here represent an important step for radio astronomy. They demonstrate that there are no unforeseen technical difficulties associated with millimetre VLBI, and that the sensitivity of existing systems is sufficient to observe objects brighter than a few Jy. This limit can soon be improved by the use of: (1) larger bandwidths now available with the Mk III system (112 MHz); (2) larger collecting areas (phased arrays using all antennas of both the Owens Valley and Hat Creek interferometers); (3) longer coherent integrations (the present limit was set by a crystal synchronizer not designed for high frequency use); and (4) lower noise temperature receivers.

Furthermore the observations are consistent with the millimetre wavelength radio structure of 3C84 being more compact

than its centimetre wavelength structure, as expected from the shape of the spectrum and the time scales of the flux density variations in these two wavelength ranges. Such an increase in compactness suggests that the millimetre wavelength observations are probing structure closer to the centre of activity, and therefore might be sensitive to features directly associated with the central engine, such as accretion disks.

These VLBI observations could not have been made without the logistical and technical support of a number of people at the Jet Propulsion Laboratory. Chief among these was Dr N. A. Renzetti, who managed to find a replacement Mark III recorder and maser at very short notice, but for which we would have been forced to abandon these observations. We thank L. J. Skjerve, R. Taylor and S. Paine, who installed these systems at Hat Creek. We are most grateful to Dr J. White, of the Naval Research Laboratory, and Dr H. Wang of the Hughes Research Laboratory, for the loan of a hydrogen maser frequency standard at short notice. We thank R. F. C. Vessot for the loan of the maser frequency standard to FCRAO. This research was supported by grants from the NSF, NASA, and the California Space Institute.

Received 13 December 1982; accepted 8 April 1983.

1. Readhead, A. C. S. *Objects of High Redshift*, IAU Symp. No. 98, 165–176 (1980).
2. Kellermann, K. I. & Pauliny-Toth, I. I. A. *Rev. Astr. Astrophys.* **19**, 373–410 (1981).
3. Readhead, A. C. S. & Pearson, T. J. *Extragalactic Radio Sources*, IAU Symp. No. 97, 279–288 (1982).
4. Cohen, M. H. *et al. Nature* **268**, 405–409 (1977).
5. Cohen, M. H. & Unwin, S. C. *Extragalactic Radio Sources*, IAU Symp. No. 97, 345–354 (1982).
6. Blandford, R. D. & Rees, M. J. *Pittsburgh Conference on BL Lac Objects*, 328–341 (1978).
7. Rogers, A. E. E. *et al. Science* **219**, 51–54 (1983).
8. Moran, J. M. *et al. Astrophys. J. Lett.* **231**, L67 (1979).
9. Moran, J. M. in *Methods of Experimental Physics*, Vol. 12, Part C (ed. Meeks, M. L.) 174–197 (Academic, New York, 1976).
10. Readhead, A. C. S., Hough, D. H., Ewing, M. S., Walker, R. C. & Romney, J. D. *Astrophys. J.* **265**, 107–131 (1983).

Reversible isothermal sorption of hydrogen by tungsten trioxide in presence of platinum

Allan R. Berzins* & Paul A. Sermon

Department of Chemistry, Brunel University, Uxbridge UB8 3PH, UK

It has long been known that WO_3 absorbs gaseous hydrogen in the presence of small quantities of a group VIII metal to form the non-stoichiometric bronze H_xWO_3 , where $0 < x < 0.5$. We now report the first demonstration that this absorption is almost entirely reversible at 273–423 K with molecular hydrogen being desorbed with decreasing pressure (0–140 kPa) at these temperatures. That desorption is preferred to reduction–dehydration to a lower oxide is unusual.

* Present address: Johnson Matthey Research Centre, Blounts Court, Sonning Common, Near Reading RG4 9NH, UK.

Non-stoichiometric tungsten bronzes A_xWO_3 , where normally $0 < x < 1.0$, are formed by simultaneous acceptance of cations A^{x+} and electrons into the interstitial channels and conduction band of WO_3 . These have long been known for tungsten¹ and subsequently bronzes of many transition metal oxides with a variety of inserted cations have been reported²⁻⁵. Hydrogen bronzes prepared electrochemically and by spillover⁶ have been reported for the oxides of V (refs 7-9), Mo (ref. 10), W (ref. 10) and Re (refs 11, 12) with maximum values of x of 3.8, 1.7, 0.5 and 0.1 respectively. Alkali metal and hydrogen bronzes of tungsten are not adequately described by the rigid band model¹³, but show similar photoelectron spectra^{14,15} with cations in similar interstitial sites and in D_xWO_3 each deuterium is bonded to the nearest oxygen in the surrounding O octahedra¹⁶. Although evidence of specific O-H bonding is restricted to inelastic neutron scattering at low temperature¹⁷⁻¹⁹ it appears that hydrogen within structurally similar ReO_3 and WO_3 is not fully protonic^{11,12,20}. Therefore a true representation of the hydrogen tungsten bronze may be between H_xWO_3 and $(OH)_xWO_{3-x}$ with only transient specific O-H interactions²⁰. This is consistent with NMR evidence of the high mobility of hydrogen within the bronze²¹⁻²⁴ but lower proton diffusion coefficients have been measured for anhydrous polycrystalline samples²⁵.

It has been inferred²⁶ that oxidation and reduction-dehydration of $H_{0.4}WO_3$ at 298 K are preferred thermodynamically to hydrogen desorption. Thus, although the sorbed hydrogen has been titrated using a variety of gaseous acceptors^{5,6,27}, heating or evacuating hydrogen bronzes of WO_3 , MoO_3 and V_2O_5 in the presence of activating metals has never desorbed more than 7% of the sorbed hydrogen in the molecular state^{6-9,28}. It is therefore surprising that recent information from temperature programmed desorption²⁹ has indicated that up to 87% of hydrogen absorbed by Pt/ WO_3 at 195 K and 101 kPa H_2 could be desorbed thermally into a N_2 stream at 273-673 K. Since H_xWO_3 prepared by an aqueous route and stored in N_2 at ambient temperature only released hydrogen after deposition of colloidal Pt, it must be assumed that the Group VIII metal accelerated the thermal desorption of hydrogen to an extent equivalent to that for the absorption process²⁹.

Samples of Pt/ WO_3 containing 0.2-5.0% Pt were prepared by impregnating batches of the trioxide (Koch Light, 99.9% purity) of differing surface area (I: $3.91 \text{ m}^2 \text{ g}^{-1}$; II: $11.15 \text{ m}^2 \text{ g}^{-1}$ estimated by N_2 BET measurements at 77 K) and average particle size (I: $8.3 \mu\text{m}$; II: 290 nm estimated by electron microscopy) with solutions of hexachloroplatinic acid (Johnson Matthey).

These were dried at 368 K, reduced in hydrogen (101 kPa, 473 K, 30 min) and reoxidized in air (573 K, 60 min). The samples were of low porosity (I: $0.17 \text{ cm}^3 \text{ g}^{-1}$; II: $0.44 \text{ cm}^3 \text{ g}^{-1}$) and their general morphology is illustrated in Fig. 1. Isothermal sorption data were corrected for the quantity of platinum-held hydrogen estimated by selective CO chemisorption. Pt dispersions decreased (and total sample surface areas increased) with increasing Pt concentration, but did not change significantly during isothermal sorption measurements. Repeated isotherms were always coincident to within $\pm 3\%$. Sorption pressures were accurate to $\pm 50 \text{ Pa}$ and values of x to ± 0.005 . In these conditions (273-498 K and 0-140 kPa H_2) weight losses were $< 0.1\%$ and it is inferred that reduction-dehydration of the bronze was insignificant. Certainly X-ray photoelectron spectroscopy indicated that very little of the tungsten was in the +5 oxidation state and differential-scanning calorimetry revealed no endotherm corresponding to dehydration of H_xWO_3 in 101 kPa H_2 below 513 K.

Figure 2 illustrates the high reversibility and reproducibility of hydrogen absorption-desorption in repeated isotherms on 3% Pt/ WO_3 (sample II) at 423 K. A comparison of the isotherms obtained using identical concentrations of Pt on different samples of trioxide in Fig. 2 reveals a marginal increase in the extent of absorption as the oxide surface area increases. This is tentatively attributed to differences in the extent of

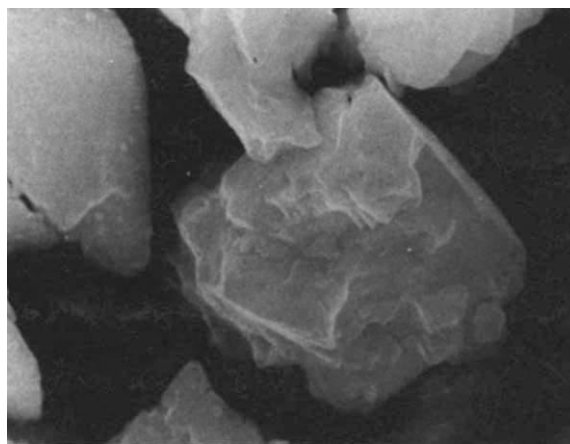


Fig. 1 Scanning electron micrograph ($\times 1,900$) of 3% Pt/ WO_3 (sample I).

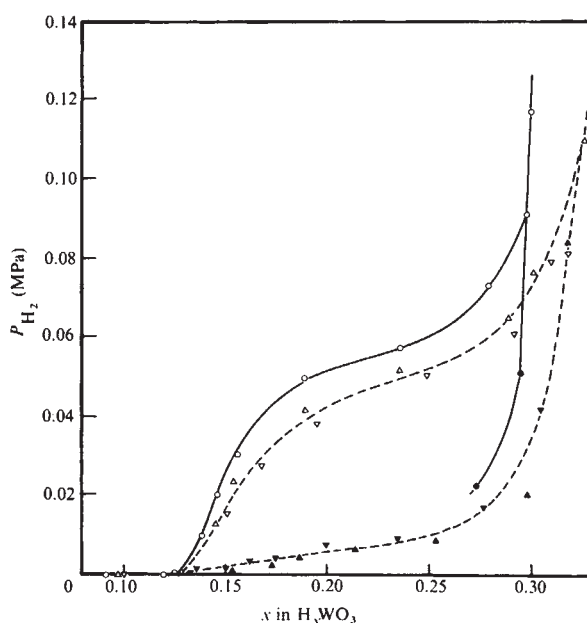


Fig. 2 Isotherms of sorption of hydrogen by 3% Pt/ WO_3 at 423 K, where the first and sixth consecutive isotherms on 3% Pt/ WO_3 (sample II) are denoted by Δ , $\blacktriangle$ and ∇ , $\blacktriangledown$ respectively and $\circ$, $\bullet$ relate to the isotherm for 3% Pt/ WO_3 (sample I). Open and closed symbols denote absorption and desorption points respectively.

absorption at low pressure where hydrogen spilt-over from the platinum is held at the surface rather than the bulk of the oxide, possibly generating anionic vacancies. Figure 3 shows that the precise concentration of Pt in the range 0.2-5.0% does not affect the extent of hydrogen sorption at 423 K by sample I of WO_3 . However, in the absence of Pt hydrogen sorption by WO_3 was not detectable at 423 K.

For hydrogen bronzes it has been suggested³⁰ that as the temperature rises specific O-H bonds are formed between lattice anions and the interstitial protons leading to production of water and a lower oxide. Therefore a significant fraction of absorbed hydrogen is not expected to be desorbed from Pt/ H_xWO_3 in the molecular state on increasing the temperature. In the light of these comments the series of hydrogen sorption isotherms at 273-498 K for 5% Pt/ WO_3 shown in Fig. 4 are interesting. Differential scanning calorimetry at 101 kPa H_2 and 423 K showed two exothermic stages in hydrogen absorption; the first with an enthalpy change of $-75.5 \pm \text{kJ mol}^{-1} H_2$ at $x < 0.12$ associated with initial low-pressure absorption, and the second with a smaller enthalpy change of $-35.5 \pm$

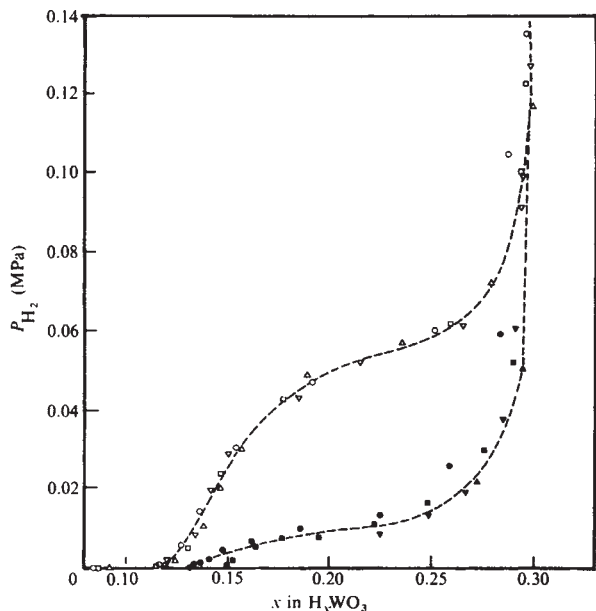


Fig. 3 Isotherms of sorption of hydrogen for Pt/WO₃ (sample I) at 423 K with different Pt concentrations (○, ●, 0.2% Pt; □, ■, 1.0% Pt; △, ▲, 3.0% Pt; ▽, ▼, 5.0% Pt). Open and closed symbols denote absorption and desorption points.

3 kJ mol⁻¹ H₂ at 0.12 < x < 0.297 associated with the absorption at higher pressure. As expected for an exothermic absorber, the extent of absorption in Fig. 4 at constant pressure decreases with increasing temperature.

The absorption-desorption hysteresis exhibited in Fig. 4 below a critical temperature of about 490 K is intriguing and has been restricted previously^{31,32} to metal hydrides where regions of change in x with little change in pressure involve transitions between two coexisting hydride phases. Equivalent plateau regions in plots of ΔG against x for H₂WO₃ may also involve the coexistence of different tetragonal phases^{33,34}. Asymmetry makes differentiation of such regions in the p-x-T plots of the Pt/WO₃-Pt/H₂WO₃ system observed here difficult. *In situ* X-ray analysis of 5% Pt/WO₃ at ambient temperature during hydrogen absorption has shown that for x < 0.03 the monoclinic structure of WO₃ is retained. When x is increased to 0.06–0.18 a tetragonal (a = 0.523 nm, c = 0.387 nm) phase is observed, which converts to a tetragonal-pseudocubic structure (a = 0.382 nm, c = 0.375 nm) when x > 0.2. Such preliminary results are consistent with earlier analyses^{14–16} but further work is required to confirm the importance of phase changes in determining the extent of isothermal hysteresis exhibited.

That the extent of desorption of hydrogen from Pt/H₂WO₃ at 423 K should so far exceed the extent of dehydration to produce reversible sorption isotherms in these conditions is

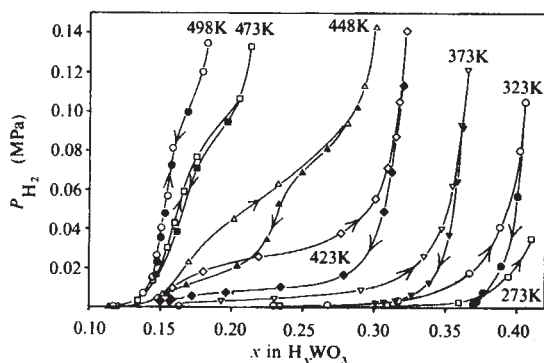


Fig. 4 Isotherms for sorption of hydrogen by 5% Pt/WO₃ at 273, 323, 373, 423, 448, 473 and 498 K, where open and closed symbols denote absorption and desorption points respectively.

surprising. Despite the proximity of H and O atoms of widely differing electronegativity, this particular bronze does form and then decompose reversibly at 273–498 K in the presence of platinum. The fraction of absorbed hydrogen recoverable at 423 K is high (up to 83%) and is not greatly dependent on the precise Pt concentration in the range 0.2–5.0 wt%. Such results have not been reported previously and the causes of hysteresis are being considered in detail and will be reported elsewhere.

The support of ARB by SERC and Johnson Matthey Co. Ltd is gratefully acknowledged.

Received 10 January; accepted 24 March 1983.

1. Wohler, F. *Ann. Phys.* **2**, 350 (1824).
2. Dickens, P. G. & Whittingham, M. S. *Q. Rev.* **22**, 30–44 (1968).
3. Hagenmuller, P. *Prog. Solid State Chem.* **5**, 71–144 (1971).
4. Bevan, D. J. M. & Hagenmuller, P. *Non-Stoichiometric Materials* (Pergamon, Oxford, 1973).
5. Schollhorn, R. *Angew. Chem.* **19**, 983–1003 (1980).
6. Sermon, P. A. & Bond, G. C. *Catal. Rev.* **8**, 211–240 (1973).
7. Schollhorn, R., Klein-Reisink, F. & Reimold, R. *JCS chem. Commun.* 398–399 (1979).
8. Tinetti, D., Estrade-Szwarcopf, H. & Fripiat, J. J. *Metal-Hydrogen Systems* (ed. Verziroglu, T. N.) 459 (Pergamon, Oxford, 1982).
9. Marq, J. P., Poncellet, G. & Fripiat, J. J. *J. Catal.* (in the press).
10. Sermon, P. A. & Bond, G. C. *JCS Faraday Trans. I* **72**, 730–744 (1976).
11. Horiuchi, S., Kimizuka, N. & Yamamoto, A. *Nature* **279**, 226–227 (1979).
12. Fujimori, A. *et al. Phys. Lett.* **80A**, 188–190 (1980).
13. Hochat, H., Bringans, R. D. & Shanks, H. R. *Solid St. Commun.* **37**, 41–44 (1980).
14. Dickens, P. G. & Hurditch, R. J. *Nature* **215**, 1266–1267 (1967).
15. Wertheim, G. K., Compagna, M., Chazalviel, J. N. & Buchanan, D. N. E. *Appl. Phys.* **13**, 225–230 (1977).
16. Wiseman, P. K. & Dickens, P. G. *J. Solid St. Chem.* **6**, 374–377 (1973).
17. Sienko, M. J. & Oesterreicher, H. *J. Am. chem. Soc.* **90**, 6568–6570 (1968).
18. Taylor, R. E., Ryan, L. M., Tindall, P. & Gerstein, B. C. *J. phys. Chem.* **73**, 5500–5507 (1969).
19. Wright, C. J. *J. Solid St. Chem.* **20**, 89–92 (1977).
20. Salje, E. & Hoppman, G. *Phil. Mag.* **43B**, 105–114 (1981).
21. Vannice, M. A., Boudart, M. & Fripiat, J. J. *J. Catal.* **17**, 359–365 (1970).
22. Nishimura, K. *Solid St. Commun.* **20**, 523–524 (1976).
23. Dickens, P. G., Murphy, D. J. & Halstead, T. K. *J. Solid St. Chem.* **6**, 370–373 (1973).
24. Hitchman, M. L. *Thin Solid Films* **61**, 341–348 (1979).
25. Randin, J. P. & Viennet, R. *J. electrochem. Soc.* **129**, 2349–2354 (1982).
26. Dickens, P. G., Moore, J. H. & Nield, D. J. *J. Solid St. Chem.* **7**, 241–244 (1973).
27. Sermon, P. A. & Bond, G. C. *JCS Faraday Trans. I* **76**, 889–900 (1980).
28. Tinetti, D. & Fripiat, J. J. *J. chim. Phys.* **76**, 867–872 (1979).
29. Sermon, P. A. & Berzins, A. R. *Metal-Hydrogen Systems* (ed. Verziroglu, T. N.) 451–457 (Pergamon, Oxford, 1982).
30. Tinetti, D. & Fripiat, J. J. *Rev. Chim. Miner.* (in the press).
31. Cohen, R. L. & Wernick, J. H. *Science* **214**, 1081–1087 (1981).
32. Lewis, F. A. *Plat. Metals Rev.* **26**, 20–27, 70–78, 121–128 (1982).
33. Dickens, P. G., Hibble, S. J. & Jarman, R. H. *J. electron. Mater.* **10**, 999–1009 (1981).
34. Jarman, R. H. & Dickens, P. G. *J. electrochem. Soc.* **129**, 2276–2278 (1982).

Dry ice II, a new polymorph of CO₂

Lin-gun Liu

Research School of Earth Sciences, Australian National University, Canberra, ACT 2600, Australia

Using a gasketed diamond-anvil press, a new polymorph of CO₂—dry ice II—has been observed in the pressure range between about 5 and 23 kbar at room temperature. The new phase has been identified by both optical and *in situ* X-ray diffraction studies which are now described. Because of the strong tendency to form large single crystals, no powder diffraction pattern for dry ice II can be obtained except in a case where a metastable implosive transition from liquid CO₂ to dry ice II occurred. The quality of the powder diffraction data thus produced, however, is too poor to deduce a possible structural type for the new phase. On the basis of optical observations, dry ice II might have a refractive index and/or density very similar to ordinary dry ice (or dry ice I) in the vicinity of 23 kbar at room temperature.

Next to H₂O, CO₂ is probably the most important gaseous species with respect to geological processes on the Earth's surface and within its interior. CO₂ is also probably an important component of other planets in our Solar System (such as Venus, the outer planets and their satellites).

Gaseous CO₂ condenses directly to solid (the ordinary dry ice, or dry ice I in this study) at –78.5°C and 1 atmospheric

pressure, and to liquid CO₂ in the vicinity of 75 bar at room temperature^{1,2}. Although numerous studies of the thermochemical properties and equations of state of gaseous and liquid CO₂ have been carried out at high pressures and temperatures¹⁻⁶, very few investigations of the solid phase at high pressure can be found in the literature.

Based on one experimental observation, Bridgman⁷ has reported that liquid CO₂ transforms to solid CO₂ at about 5.2 kbar and room temperature; he⁸ also reported that a solid-solid transition may occur at pressures of 25 kbar or higher (at the solid CO₂ temperature?, -78.5 °C). Hanson *et al.*⁹ carried out a Raman study of solid CO₂ to 110 kbar at room temperature. However, it is not known what is the solid phase in equilibrium with liquid CO₂ at high pressure and room temperature, which is always implicitly assumed to be dry ice I although it has never been confirmed experimentally. Dry ice I has also been studied by shock-wave experiments to about 670 kbar (ref. 10), but again the state of CO₂ at high pressures is unknown.

In the present investigation, I performed *in situ* studies of high-pressure phase transformations of CO₂ by means of both optical and X-ray observations, in a standard gasketed diamond-anvil pressure cell. The cell was cooled to the dry ice temperature and/or liquid nitrogen temperature, then a sample of solid CO₂ was loaded into the gasketed area of the cooled cell which was then sealed immediately at these temperatures. Both commercial dry ice and solid CO₂ made from a commercial grade CO₂ gas cylinder in our laboratory were used in the study. In most cases, samples were loaded in the ambient air atmosphere. Thus, moisture from air was also inevitably included as an impurity. However, one sample was loaded in a CO₂ atmosphere for cross-checking. Optical and X-ray studies were subsequently done at room temperature. Pressures were measured using the ruby-fluorescence technique¹¹.

As mentioned above, there is only one experimental observation that indicates that liquid CO₂ transforms to solid CO₂ at about 5.2 kbar and 25 °C. In the present study, from five sets of experiments using two different diamond-anvil cells nine attempts were made to observe the coexistence of liquid and solid CO₂ (compare with my earlier study¹² of H₂O), but all experiments have so far failed. In the increasing pressure runs, the lowest pressure at which the solid phase appears suddenly is 8.7 kbar. In the decreasing pressure runs, the lowest pressure at which CO₂ remains solid is 5.4 kbar, generally confirming the results of Bridgman⁷. However, in the present study a metastability in the liquid to solid transition in CO₂ apparently exists at room temperature. A supercooled liquid can exist in

Table 1 Powder diffraction pattern for the new polymorph CO₂ at 17.1 kbar and room temperature (Mo K α radiation)

Intensity	d-spacing (Å)
Medium	2.76
Strong	2.626
Weak	2.04
Medium	1.93
Weak	1.83
Strong	1.61

a solid stability field when temperature of a liquid phase decreases to a solid phase. Superpressed liquids can also occur when a liquid is solidified by applying pressure.

I was unable to obtain an X-ray powder diffraction pattern for the solid phase crystallized from liquid CO₂ because of the large single crystals formed in all of my experiments except for the one case discussed below. Regardless of the different samples used and the different loading environments, the solid phase was further observed to transform to another solid phase at 23 $\pm$ 1 kbar and room temperature; this transition was sometimes optically observed. The coexistence of the two phases was not as clear as in the case of my previous study of the ice VI-ice VII transition, suggesting that the two phases have very similar refractive indices and/or densities. *In situ* X-ray diffraction studies revealed that the second solid phase is dry ice I in all cases. A reversed experiment was also carried out to check the reality of the new solid phase between liquid CO₂ and dry ice I at room temperature. Dry ice I at high pressures (as confirmed by X-ray diffraction) was brought to a pressure of $\sim$ 20 kbar, and an X-ray diffraction study showed that the powder diffraction pattern of dry ice I was replaced by single-crystal diffraction spots of the new phase. This confirms the existence of the new solid phase between liquid CO₂ and dry ice I at room temperature. A tentative pressure-temperature phase diagram for CO₂ based on the results of this study and those reported in the literature^{1,2,7,13} is presented in Fig. 1.

The only powder diffraction pattern for the new polymorph of CO₂ was obtained when a 'metastable' liquid CO₂ suddenly transformed to the new solid phase, accompanied by an implosion. In one of my attempts to obtain coexisting liquid and solid CO₂, liquid CO₂ was found to sustain to 13.2 kbar, which is more than twice the pressure of the supposed boundary reported by Bridgman⁷. When more pressure was applied, a loud implosion was audible and a phase change was observed optically. After the implosion, the sample changed from colourless to pale-yellow and the pressure was 17.1 kbar. X-ray diffraction of the sample showed powder diffraction lines of rather poor quality; the intensities and d-spacings of these are given in Table 1. No attempt has been made to deduce a possible structural type for the new polymorph based on the very limited data of the present study. However, it seems certain that these diffraction lines do not correspond to the pattern of dry ice I. The same diffraction pattern was observed at 20.9 kbar, but it transforms to dry ice I at pressures greater than 23 kbar. Following the terminology used for H₂O, the new polymorph of CO₂ can be termed 'dry ice II' as shown in Fig. 1.

I thank R. D. Holmes for helpful comments.

Received 11 February; accepted 28 March 1983.

- Villard, P. & Jarry, R. *C. r. hebdomadaire Séances Acad. Sci. Paris* **120**, 1413-1416 (1895).
- Michels, A. & Michels, C. *Proc. R. Soc. A* **153**, 201-214 (1935).
- Michels, A., Michels, C. & Wouters, H. *Proc. R. Soc. A* **153**, 214-224 (1935).
- Kennedy, G. C. *Am. J. Sci.* **252**, 225-241 (1954).
- Altunin, V. V. & Gadetskii, O. G. *Teploenergetika* **18**, 120-125 (1971).
- Shmonov, V. M. & Shmulovich, K. I. *Dokl. Akad. Nauk. SSSR, Earth Sci.* **217**, 206-209 (1975).
- Bridgman, P. W. *Phys. Rev.* **3**, 153-203 (1914).
- Bridgman, P. W. *Proc. Am. Acad. Arts Sci.* **72**, 207-225 (1938).
- Hanson, R. C., Jordan, M. & Bachman, K. in *High Pressure Science and Technology* (eds Vodar, B. & Marteau, P.) 763-765 (Pergamon, Oxford, 1980).
- Zubarev, V. N. & Telegin, G. S. *Soviet Phys.-Dokl.* **7**, 34-36 (1962).
- Barnett, J. D., Block, S. & Piermarini, G. J. *Rev. scient. Instrum.* **44**, 1-9 (1973).
- Liu, L. *Earth planet. Sci. Lett.* **61**, 359-364 (1982).
- Michels, A., Blaise, B. & Hoogschagen, J. *Physica* **9**, 565-573 (1942).

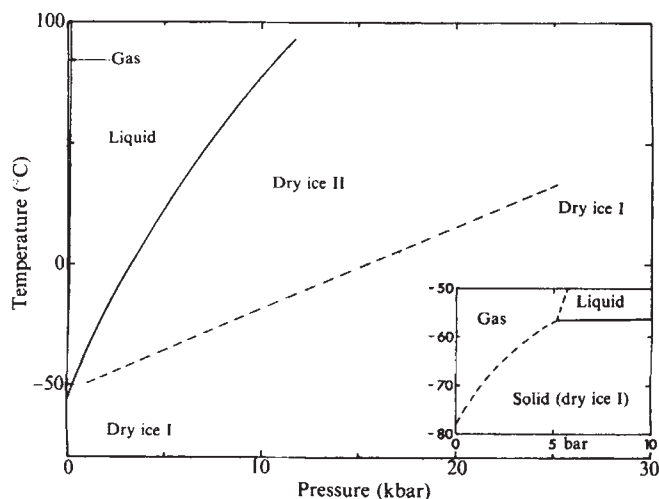


Fig. 1 A tentative *P-T* phase diagram for CO₂ based on the results of this study and those reported in the literature^{1,2,7,13}. The suggested small *dT/dP* slope for the phase boundary between dry ice II and dry ice I is consistent with the small density difference between the two phases observed in the present study.

Interpretation of sediment cores from the Ross Ice Shelf Site J-9, Antarctica

Howard Thomas Brady

Macquarie University, Sydney and Hordern Utz and Bode,
15-19 Bent Street, Sydney, Australia 2000

Stratigraphical interpretations of sediments beneath the Ross Ice Shelf, Antarctica, are derived from 58 short gravity cores (up to 1.22 m in length) taken by Webb and Brady through an ice access hole at the Ross Ice Shelf (RISP) Site J-9 (82°22' S), 168°38' W) during the austral summers of 1977 and 1978. Some scientific reports on microfossil groups of diatoms^{1,2}, foraminifera³, and silicoflagellates⁴ have interpreted the sedimentary succession as *in situ* and of Miocene age. Webb⁵ has also reported reworked clasts of Cretaceous foraminifera. Brady and Martin² suggested that the pollen assemblage within the cores, similar to a reported Oligocene flora⁶, may have reflected a remnant of this flora living in the Miocene, but Wrenn⁷ has shown that these pollen have probably been reworked. However, Kellogg and Kellogg's⁸ diatom-based Pleistocene age determinations for sediment exposed on the sea floor at Site J9 relies on misquoted data from published literature and on misinterpreted taxonomy. Evidence is presented here that supports an earlier proposal that Miocene sediments crop out on the present sea floor.

Kellogg and Kellogg's interpretation⁸ stated that the J-9 sedimentary succession was Recent throughout because the cores contained Pleistocene, Pliocene and Miocene diatoms. These authors published a stratigraphical range table of fossil diatoms in the J-9 cores of which five were placed in the Pleistocene, two within the Plio-Pleistocene, and one within the mid-Pliocene. This table indicated two other diatoms which may range beyond the Pleistocene. In a separate distribution table the authors listed another Plio-Pleistocene diatom—*Thalassiosira oestrupii*.

Unfortunately, the stratigraphical range table published by Kellogg and Kellogg which they have derived from the DSDP reports of McCollum⁹ (Leg 28) and Schrader¹⁰ (Legs 29 and 35) misrepresents the views of these authors on at least eight occasions. It is also possible that other Plio-Pleistocene taxa (especially *Nitzschia curta*) may have been misidentified and are not present in the J-9 cores.

Some of the misunderstandings in the Kelloggs' report may arise from the fact that some of the diatom zones published by McCollum⁹ in the DSDP Leg 28 report are clearly acme zones (zones of abundance) and not time zones in the strict sense based on the last and first appearance of a particular taxon. This becomes clear on inspection of the distribution charts published by McCollum for particular DSDP holes (for example, McCollum⁹ records the appearance of *Eucampia balaustium-antarctica* from the middle and late Miocene but his abundance zone for this taxon shows a Plio-Pleistocene range). It is more difficult, however, to see how Kellogg and Kellogg⁸ misunderstood the diatom report of Schrader¹⁰ (DSDP 29 and 35) as in some cases both his distribution charts and his time range zones are misquoted. Furthermore, the report of Fenner quoted by Kellogg and Kellogg⁸ does not exist as quoted.

Because at the recent geological meeting of the Scientific Committee for Antarctic Research in Adelaide (August 1982), discussions indicated there would be no drilling into sediments beneath the Ross Ice Shelf in the near future, it is particularly necessary to evaluate the fossil evidence presented by Kellogg and Kellogg⁸. I therefore now discuss¹¹ taxa whose stratigraphical ranges or identification are questionable:

Coscinodiscus stellaris var. *symbolophorus*: quoted as Brun-

hes in age according to McCollum⁹ by Kellogg and Kellogg⁸, this taxon is actually recorded by McCollum in the middle Miocene of DSDP 273 (core 6-1, p. 527 of ref. 9) and by Schrader throughout the whole Miocene of DSDP 278 (cores 10-2 to 30-1, pp. 622-624 of ref. 10). Furthermore, a question mark beneath the zonal range bar of the taxon in McCollum's diagram (p. 532 of ref. 9) was omitted in the Brunhes zonal range bar of Kellogg and Kellogg (Fig. 1 of ref. 8).

Eucampia balaustium-antarctica: quoting McCollum⁹, Kellogg and Kellogg⁸ describe this taxon as Plio-Pleistocene in age. However, McCollum records it in seven samples from the late-middle Miocene and late Miocene in DSDP 269 (cores 3-3 to 6-3, p. 526 of ref. 9). In addition, no reference is made to the work of Schrader who records this taxon in 53 samples from the middle and late Miocene of DSDP 278 (cores 10-22, p. 620 of ref. 10).

Rouxia antarctica: McCollum (p. 532 of ref. 9) indicates that this taxon has abundance zones in the Pleistocene and Pliocene. Question marks indicate that he did not know the time range for the taxon. Unfortunately, Kellogg and Kellogg⁸ omitted to report McCollum's Pliocene abundance zone with the question mark which indicated that he thought its first appearance was earlier and they only record the abundance zone and associated question mark in the Pleistocene. They also failed to note Schrader's record of the taxon from the late to middle Miocene of DSDP 278 (cores 16, 17 and perhaps 21?, pp. 626-627 of ref. 10).

Asteromphalus parvulus: Kellogg and Kellogg⁸ rightly note that McCollum records this taxon from the Brunhes of DSDP 266 and from a mixed Pleistocene-Miocene interval of DSDP 273. However, Schrader recorded the taxon from the early Miocene of DSDP 278 (cores 27-1 and 29-4, pp. 623-624 of ref. 10). Not enough is known about this taxon to use it as a stratigraphical marker species.

Thalassionema elegans: quoting McCollum⁹, Kellogg and Kellogg⁸ describe this taxon as Brunhes, yet McCollum records it from the early Pliocene or late Miocene of DSDP 274 (cores 10-2, 10-4, 10-6, p. 529 of ref. 9). Since it is recorded by McCollum in only one DSDP hole, this taxon cannot be regarded as an established stratigraphical marker species (even if it has been correctly identified by Kellogg and Kellogg in J-9 cores).

Schimperella antarctica: quoting Schrader¹⁰, Kellogg and Kellogg⁸ describe this taxon as Brunhes in age. Schrader records the species from the mid-Pliocene of DSDP 322 (core 1-1, p. 609 of ref. 10) and from the early Pliocene of DSDP 323 (core 2-1, p. 611 of ref. 10). This taxon was not recorded by McCollum⁹ during DSDP Leg 28. A reference to Fenner, DSDP Leg 39, does not apparently exist and the occurrence of the species in a plankton survey during DSDP Leg 35 by Fenner *et al.*¹¹ has no biostratigraphical significance.

Charotia actinochilus: Kellogg and Kellogg⁸ record this species as Pleistocene according to both Schrader and Fenner and a question mark in the Kellogg's diagram indicates a possibly earlier age. McCollum⁹ records the species from the Pleistocene of DSDP 269, 272, 273 and 274 and Schrader¹⁰ in the Pleistocene of DSDP 323 and 324. More study may indicate that this species may be restricted to the Pleistocene but the species may easily be confused with various forms of *Actinocyclus* and *Cestodiscus*.

Thalassiosira gracilis: Kellogg and Kellogg⁸ correctly record McCollum's Plio-Pleistocene range for this taxon. In one sample from J-9 core no. 7 they note that this taxon is 15% of the flora. This author re-examined nine samples from J-9 core 7 and did not observe this taxon.

Nitzschia curta: Kellogg and Kellogg⁸ counted 130 individuals of this taxon in a 4,500 diatom count of 14 samples from J-9 cores 7 and 8. Because they reported this taxon as common, I have scanned two samples from J-9 core 7 and taken scanning electron microscope photographs of *Nitzschia* that could resemble *Nitzschia curta*. One taxon (called *Nitzschia truncata* n. sp. by Brady¹) accounted for 2-7% of the flora in

6–9 samples studied by me. This taxon may be confused with *Nitzschia curta* under the light microscope. Scanning electron microscopy studies show that its striae density and the arrangement of pores within the striae are completely different from *Nitzschia curta* (24 striae per 10 μm as against 14–17 striae per 10 μm , and striae with single rows of pores as against occasional sections of striae with double rows in *Nitzschia curta*). Both species have short curved rows of apical striae at right angles to the transverse rows of striae elsewhere in the valve, but finer details are completely different).

Nitzschia interfrigidaria: this is correctly reported by Kellogg and Kellogg⁸ as a mid-Pliocene marker species. I suspect that they may have confused this species with *Nitzschia maleinterpretaria* Schrader or *Nitzschia* no. 17 Schrader which Schrader¹⁰ recorded in the Miocene of DSDP 278 and which I have noted in the J-9 cores and in the middle Miocene of DSDP 266.

Thalassiosira oestrupi: Kellogg and Kellogg⁸ note this species but do not comment on its stratigraphical range. I have noted it in the Pleistocene to early Pliocene in Leg 28 cores. The species is noted by an offset pore in the centre of the valve and by one offset pore three areolae out from the centre of the valve. I have noted one species of *Thalassiosira* in J-9 cores with an offset centre pore and a series of three separate pores spaced three areolae out from the centre of the valve and I suspect that this taxon was mistaken for *Thalassiosira oestrupi* by Kellogg and Kellogg⁸.

Apart from their biostratigraphy and taxonomy, other statements made by Kellogg and Kellogg⁸ may be queried. Brady and Martin² did not use pollen and spores to support a Miocene age for the J-9 cores. Rather they proposed a hypothesis that the pollen flora in the J-9 cores (similar to the late Oligocene assemblage from the Ross Sea) may not be reworked but may represent a surviving remnant of the earlier Oligocene forests. Dating of the J-9 cores relied on diatoms, foraminifera and silicoflagellates, not on pollen. Kellogg and Kellogg⁸ also state that neither Schrader nor Brady disagree with McCollum's Pleistocene diatom biostratigraphy. Schrader¹⁰ has revised McCollum's diatom biostratigraphy for the Miocene using DSDP 278 (Southern Tasman Sea) but he did not revise McCollum's zonations for the Plio-Pleistocene because as shipboard scientist on DSDP Leg 35 (Bellinghausen Sea) Schrader found the Plio-Pleistocene section poorly preserved. As a consequence and for the purpose of the Leg 35 report Schrader simply used McCollum's zones unchanged. I have re-examined all the Leg 28 material used by McCollum in his initial Leg 28 report and acknowledge McCollum's pioneering work. However, even in the Plio-Pleistocene (as in the Miocene) McCollum's work does require refinement so that true stratigraphical marker species are clearly defined and separated not only from rare species whose stratigraphical range is unclear but also from other long ranging Miocene to Recent species with abundance zones in the Plio-Pleistocene. Furthermore, McCollum overlooked many small species of *Nitzschia*—especially in the Miocene.

In my view, the problems of the J-9 cores and their interpretation in relation to ice levels in the Ross Sea embayment have been obscured in the paper by Kellogg and Kellogg⁸. It is clear from the DSDP Leg 28 reports that there are widespread unconformities related to grounding ice sheets throughout Ross Sea sediments.

At DSDP 270, the angular unconformity apparent on the seismic profile at 20 m was interpreted as a boundary between softer early Pliocene sediments and indurated early Miocene sediments¹². At DSDP 272, a similar unconformity was recognized 25 m downhole between early Pliocene and late Miocene (p. 212 of ref. 12) or middle Miocene (p. 8 of ref. 12) sediments. And in DSDP 273, an unconformity was noted at 42 m between Recent and middle Miocene sediments (p. 341 of ref. 12). Such unconformities may be related either to the advance or the retreat phases of grounded ice masses. In some places the erosion of older sediments and the deposition of new mixed sediments could be simultaneous. However, the existence of

such unconformities open the possibility that in some places (particularly in the Southern Ross Sea) there may have been erosion and no subsequent deposition of sediments so that old sediments are exposed on the sea floor. I maintain that such is the case at the RISP J-9 site and that Kellogg and Kellogg⁸ have not established the existence of sediments younger than Miocene on the sea floor at the RISP J-9 site.

I thank Horden Utz and Bode (Member Firm of the Sydney Stock Exchange) and the School of Biological Sciences, Macquarie University for financially supporting this work.

Received 8 October 1982; accepted 30 March 1983.

1. Brady, H. T. *Antarct. J. U.S.A.* **13**, 123–4 (1978).
2. Brady, H. T. & Martin, H. *Science* **203**, 437–8 (1979).
3. Webb, P. N. *RISP Tech. Rep.* 79-1 (University of Nebraska, 1979).
4. Ling, H. Y. *4th Symp. on Antarctic Geology and Geophysics*, University of Adelaide (in the press).
5. Webb, P. N. *J. R. Soc. N.Z.* **11**, 439–446 (1981).
6. Barrett, P. J. & Kemp, E. M. *Nature* **258**, 507 (1975).
7. Wrenn, J. H. & Beckman, S. W. *Nature* **216**, 187–189 (1982).
8. Kellogg, T. B. & Kellogg, D. E. *Nature* **293**, 130–133 (1981).
9. McCollum, D. W. *Init. Rep. DSDP Leg 28*, 515–571 (1975).
10. Schrader, H. J. *Init. Rep. DSDP Leg 35*, 605–671 (1976).
11. Fenner, J., Schrader, H. J. & Wienig, H. *Init. Rep. DSDP Leg 35*, 757–813 (1976).
12. Hayes, D. E. *et al. Init. Rep. DSDP Leg 28*, 1–1017 (1975).

Reply by Thomas B. Kellogg & Davida E. Kellogg

Department of Geological Sciences and Institute for Quaternary Studies, University of Maine at Orono, Orono, Maine 04469, USA

DESPITE never having examined our Ross Ice Shelf Project (RISP) slides or photomicrographs, Brady cites 11 specific instances in which he alleges that we either misidentified common Plio-Pleistocene indicator species, or misrepresented their stratigraphical ranges¹. Each allegation is vital to Brady's argument that the RISP sediments are *in situ* middle Miocene in age because if any one of the indicator species we reported is present in the RISP cores and has a post-Miocene age, then his age interpretation cannot be correct. He alleges or implies that we misidentified *Thalassionema elegans*, *Charcotia actinochilus*, *Nitzschia curta*, *Nitzschia interfrigidaria*, *Thalassiosira gracilis*, and *Thalassiosira oestrupi*. It is the recognition of these post-Miocene species, none of which was reported by Brady and Martin², that led to our interpretation of RISP sediments that is the basis of the present controversy. Our samples contain a hash of diatom fragments, many of which are too small for positive identification. Identifiable specimens of all species are rare, and are dominated by the robust Miocene species reported by Brady and Martin². Against such a background, unequivocal documentation of the generally rarer, smaller, and more delicate post-Miocene species required exhaustive study, examining multiple slides of each sample. Because of the unusual difficulties presented by the reworked nature of the RISP samples, and the previously published Miocene age² for them, we took special care to key our identifications to original species' descriptions, and sought outside expert confirmation of any identifications which remained uncertain. No species was reported unless it was positively identified in at least two samples (the sole exception being possible *Charcotia actinochilus*, fragments of which always were too small for positive identification). The results of our study are being prepared for a larger publication, which will include photomicrographs of the species Brady alleges were misidentified as well as the full taxonomic and stratigraphical discussions which space limitations precluded from our earlier publication. We stand by the species identifications that we reported¹.

A case in point is the identification of the Brunhes indicator species *Nitzschia curta*, which is the most abundant of the Pleistocene species we reported¹, and which is the species to which Brady devotes the most attention in his comments. Brady

implies that we confused *N. curta* with *Nitzschia truncata*, a species he claims to have named in 1978. However, if one refers to Brady's (1978) paper³, the only mention of this species is: "Floras are dominated by *Melosira sulcata*, ... and *Nitzschia* n. sp." There is no mention of the species name *truncata*, no description, and no established type specimen. This hardly constitutes an adequate (or citable) description. The first descriptive material we have seen that apparently corresponds to *Nitzschia* n. sp.³ appears in Brady's comment (and even this is insufficient to satisfy the requirements established in the International Code of Botanical Nomenclature⁴). The name *Nitzschia truncata* first appeared in the literature in Brady's paper⁵ (where it was listed in a table, again without descriptive material). This name is invalid because in effect it constitutes a *nomen nudum* (or, quite possibly, a junior synonym for *N. curta*).

When first examining the RISP cores, we discovered small specimens that appeared similar to *N. curta* and fell in the lower part of the normal size range for this species. Hasle⁶ gave dimensions of 10–42 µm in length, 3.5–6 µm in width and 9–12 costae in 10 µm for *N. curta*. Our specimens range from 11 to >22 µm in length and 3–5 µm in width, but they have between 12 and 17 costae in 10 µm. Brady states that his specimens have 24 striae in 10 µm as against 14–17 for *N. curta*. The source of this costae count for *N. curta* is not given, although Fenner *et al.*⁷ gave a range of 11–16. Our specimens fall within the lower part of the range of published lengths for *N. curta*, have appropriate widths, and sometimes slightly exceed published numbers of costae per 10 µm. In no case do they have nearly as many costae as Brady reports. We thus see no reason to believe we have confused *N. curta* with Brady's invalid taxon. Hasle⁶ also stated that *N. curta* "... is characterized by the curved transapical costae near the poles and by the heteropole apical axis, one pole being broader than the other." Most of our specimens which are not obscured by debris and which do not have broken ends fulfill these criteria. We conclude that our specimens are *N. curta*, and we have been encouraged in this belief by independent experts who have examined our slides (for example, L. Burckle, personal communication).

What then is the significance of Brady's specimens of *Nitzschia*? These specimens are purported to have similar dimensions to *N. curta* as well as curved costae near the apices and a single row of pores between the striae. Brady claims that this single row of pores distinguishes his specimens from *N. curta*, but Hasle⁶ reported that although *N. curta* usually has two rows of poroids, it occasionally has one or even three rows. As a working hypothesis, it is possible that Brady's form is an ecophenotype of *N. curta*. The larger number of costae and smaller number of rows of poroids may occur in some specimens living near the extreme end of the range of environmental conditions tolerable by *N. curta*. If this is so, we believe that it is Brady who has misidentified *N. curta* (as *N. truncata* nom. nudum), and it is probably one of the principal reasons for his continued resistance to accepting our interpretation of the RISP cores.

More puzzling to us is Brady's allegation that we must have misidentified *Thalassionema elegans* and *Thalassiosira gracilis* and that we confused *Nitzschia interfrigidaria* with *N. grossepunctata*. These are all such distinctive species that it would be difficult to mistake them, even when only preserved as large fragments as is commonly the case in the RISP samples. We followed procedures similar to those described for *N. curta* when identifying these species.

Because space limitations precluded our discussing adequately the diatom biostratigraphical information we used, we take this opportunity to elaborate on our biostratigraphical rationale before answering Brady's specific comments.

In 1979–81, when we did our analyses of the RISP cores, several local biostratigraphical studies were available for Antarctic and Sub-Antarctic areas. These included Schrader's study of DSDP cores from the south-east Pacific (and one core from south of New Zealand⁸), Fenner's work with South

Atlantic DSDP cores⁹, and McCollum's study of Ross Sea DSDP material¹⁰. There was no synthesis available of these (and other) regional studies, as there is now¹¹. We evaluated each biostratigraphical study to determine if, and to what extent, we should rely on published stratigraphical ranges for our study of the RISP diatoms. In many cases we went beyond these stratigraphical studies to the initial core reports to determine the suitability of particular cores for providing reliable stratigraphical information. We finally decided to rely almost entirely on McCollum's work¹⁰, the only local diatom biostratigraphy for the Ross Sea.

Several factors dictated our choice of McCollum's biostratigraphy. Schrader repeatedly warned that many species in his cores were reworked by bottom currents⁸, thus casting some doubt on the reliability of his species ranges. More important, he also cautioned that many diatom species may have different stratigraphical ranges in different localities when he stated that, "It seems that at present a standard diatom stratigraphy is still not possible because of local ranges of high diversity, high ecological variation, and palaeo-geographical distribution." We therefore decided not to use Schrader's biostratigraphy, or that of Fenner⁹ except in cases where McCollum's stratigraphy was not adequate (for example, we used Fenner's work for the range of *Schimperella antarctica* because this species was not mentioned by McCollum). In short, we considered it speculative to use ranges derived from distant areas in our study of the RISP cores. Our nearly total dependence on McCollum¹⁰ for stratigraphical information is stated in the caption to our Fig. 1. The only two species whose ranges were taken from other authors were *Charcotia actinochilus* and *Schimperella antarctica*, and we quoted Schrader⁸ and Fenner⁹ for these two ranges, respectively.

We concur with Brady that McCollum's stratigraphy was a valuable pioneering effort, but we recognized some limitations to the species ranges presented there. First, although recognizing that reworking was sometimes a problem in Ross Sea DSDP cores (p. 521, of ref. 10), McCollum was not aware of our work documenting the widespread action of grounded ice¹². We suspect that he presented common-to-abundant species ranges instead of true stratigraphical ranges because he recognized that rare species occurrences might represent reworking. McCollum's Fig. 10⁸ represents his own best estimate of ranges for the species he encountered, and this was the source we used to obtain ranges for RISP diatoms. We did make some modifications. First, because we were concerned with the possibility that some species may have different local ranges, we decided to rely most heavily on those of McCollum's ranges that were based on Ross Sea material. For this reason we omitted the Miocene portion of McCollum's range for *Eucampia antarctica* because this age was based on rare occurrences in Site 269, outside the immediate Ross Sea region. Second, the Shipboard Scientific Party that described Site 274 stated that¹³ "Deformation due to drilling is unusually intense throughout most of the hole, making recognition of stratification and sedimentary structures difficult." We felt that reliance on such a badly distributed core for stratigraphical information was unwise, and therefore we omitted parts of the ranges of *Rouxia antarctica* and *Thalassionema elegans* that were based on occurrences in Site 274.

We now turn to Brady's specific comments:

We fail to understand how Brady can charge that we misrepresented Schrader's stratigraphical ranges⁸ for *Coscinodiscus stellaris* var. *symbolophorus*, *Eucampia antarctica*, *Asteromphalus parvulus*, and *Schimperella antarctica*. We could not misrepresent Schrader because we did not quote him for these species.

Coscinodiscus stellaris var. *symbolophorus*: An additional reason for not quoting Schrader for this species is that he did not report var. *symbolophorus* but on the *Coscinodiscus symbolophorus* group (p. 631 of ref. 8). Schrader either could not or did not recognize var. *symbolophorus*, and thus his stratigraphical range for the entire group is not applicable for our

work. Brady asserts that we omitted the Miocene portion of the range for this species given by McCollum. We overlooked McCollum's single occurrence of *C. stellaris* var. *symbolophorus* in site 273A (6A-1) and should have included a "?" in our Fig. 1.

Schimperella antarctica: Because McCollum¹⁰ did not report this species from the Ross Sea, we were forced to look farther afield for stratigraphical information. We did not use Schrader's range because it was based on DSDP sites 322 and 323. Schrader mentioned problems with contamination and displaced diatoms in both these cores. Fenner⁹ (not Fenner *et al.*⁷) showed a stratigraphical range in the Brunhes. Brady seems unaware of this important paper on DSDP Leg 39. It appeared in the supplementary volume (where, inside the cover page, a statement occurs recommending that references to papers in the supplementary volume be referred to the original volumes—a practice we followed). We are surprised that Brady did not challenge our identification of *S. antarctica*, because this species has a range that is clearly post-Miocene.

We recognized that many of the stratigraphical ranges we used were preliminary and local. Some ranges have already undergone revision^{11,14}. Other species may not prove to be useful stratigraphical indicators (for example, *Asteromphalus parvulus*). The major point of our paper was that the published late-middle Miocene age for the RISP sediments² is incorrect because these cores also contain Pliocene and Pleistocene diatom species. We did not state that the "... J-9 sedimentary succession was Recent throughout ..." as claimed by Brady, and we fail to understand how a diatom assemblage with this mix of stratigraphical ranges can be misconstrued as Recent (with a capital R). We used the best available local information to show that the RISP sediments were reworked, containing both Miocene and Plio-Pleistocene species.

In the light of our data, evidence that the RISP sediments are *in situ* Miocene are conspicuously lacking, and the burden of proof for this age remains on Brady. Such proof must include documentation that all the Plio-Pleistocene species we reported are either misidentifications, or that they have ranges extending into the Miocene in the Ross Sea. It is not sufficient to document a longer range in cores outside the local region (for example, Schrader's reported Miocene occurrences⁸), as Brady has suggested here. We regard our hypothesis as proved by the sole presence of *Nitzschia curta* in the RISP cores. Brady does not challenge the Pleistocene range of this species, and we regard his assertion of misidentification as unfounded, as described above. Other species we reported simply strengthen our case. In this regard, we note that Brady has not challenged our identification of *Schimperella antarctica*, but only its stratigraphical range which he gives as Pliocene to Recent. He gives a similar range for *Thalassiosira oestrupi*, but challenges our identification.

Brady also alleges that we misquoted him because, "Brady and Martin did not use pollen and spores to support a Miocene age for the J-9 cores." Yet Brady and Martin² state: "The pollen assemblage in these sediment cores is wholly consistent with the middle Miocene age of the diatoms." If this statement is not an unqualified support for the diatom-derived Miocene age, what is it?

Brady takes exception to our statement that, "... neither we nor Brady, or Schrader doubt this..." (the validity of McCollum's stratigraphical ranges for the Pliocene and Pleistocene). Here we were remiss in our choice of words. We should have stated that neither we nor Brady, or Schrader had published doubts.

We regard Brady's criticisms as unfounded. In our view, the problems of the J-9 cores have been largely solved by our own study¹, by the recent work of Faure and Taylor¹⁵ and Anderson *et al.*¹⁶, and by this opportunity to expand on some of the ideas we presented earlier. The RISP sediments are not *in situ* Miocene. They were repeatedly reworked, probably by the action of grounded ice, and the most recent episode of reworking probably occurred in the late Brunhes (21,000–

17,000 yr BP?). The main remaining problem is the precise age of the most recent episode of reworking. We hope Brady will join us in working towards a resolution of this geological problem.

Received 25 March; accepted 30 March 1983.

1. Kellogg, T. B. & Kellogg, D. E. *Nature* **293**, 130–133 (1981).
2. Brady, H. T. & Martin, H. *Science* **203**, 437–438 (1979).
3. Brady, H. T. *Antarct. J. U.S.A.* **13**, 123–124 (1978).
4. *International Code of Botanical Nomenclature* (Adopted by 11th Int. Botanical Congress, Seattle, 1969).
5. Brady, H. T. *Antarct. J. U.S.A.* **14**, 130 (1979).
6. Hasle, G. R. *Norske Videnskaps-Akademi Oslo Mat-Naturv. Klasse. Ny Ser. No. 21*, 1–49 (1965).
7. Fenner, J., Schrader, H.-J. & Wienigk, H. *Init. Rep. DSDP Leg 35*, 757–813 (1976).
8. Schrader, H.-J. *Init. Rep. DSDP Leg 35*, 605–671 (1976).
9. Fenner, J. *Init. Rep. DSDP Leg 39*, 491–623 (in Suppl. Vol. 1978).
10. McCollum, D. W. *Init. Rep. DSDP Leg 28*, 515–571 (1975).
11. Weaver, F. M. & Gombos, A. M. Jr *Soc. Econ. Paleont. Miner. Spec. Publ. No. 32*, 445–470 (1981).
12. Kellogg, T. B., Truesdale, R. S. & Osterman, L. E. *Geology* **7**, 249–253 (1979).
13. Hayes, D. E. *et al. Init. Rep. DSDP Leg 28*, 369–433 (1975).
14. Ciesielski, P. F. *Init. Rep. DSDP Leg 71* (in the press).
15. Faure, G. & Taylor, K. S. *Proc. 4th int. Symp. on Antarctic Earth Sciences, Adelaide Univ.* (in the press).
16. Anderson, J. B., Kurtz, D. D., Domack, E. W. & Balshaw, K. M. *J. Geol.* **88**, 399–414 (1980).

Palaeocontinental configurations and geoid anomalies

Poorna C. Pal

Department of Geology and Mineral Sciences, University of Ilorin, P.M.B. 1515, Ilorin, Nigeria

Reconstructions of past continental configurations help our understanding of the evolution of planetary surface morphology and impose valuable constraints on the models for planetary dynamics. The problem is that although well established palaeogeographical maps now exist for Mesozoic and later times^{1,2}, thanks to excellent agreement between palaeomagnetic and marine-magnetic deductions, this is not true for earlier times mainly due to the lack of independent palaeogeophysical control over otherwise longitudinally-ambiguous palaeomagnetic data. The prospect of relating palaeocontinental configurations to geoid anomalies, recently raised by Anderson³, has thus given a fresh perspective to this problem. Here his continental insulation model is shown to be an inadequate explanation of the undulations of subduction-compensated geoid. Instead, the geoid highs and lows are shown to correlate well with the overall divergent and convergent plate tectonism since the middle Palaeozoic and are thus ascribed, respectively, to the ascending and descending patterns of lower-mantle convection.

As for the geoid, Chase's⁴ analysis of GEM-8⁵ has revealed three enormous anomalies—two highs encompassing Africa–Atlantic and Pacific regions and a low stretching discontinuously from Antarctica through India and Siberia—after compensation for the effect of subduction. These anomalies seem totally unrelated to the present tectonic framework except that hot-spots, mostly oceanic, cluster about the 20–40 m positive geoid contours⁶. Anderson³ has therefore hypothesized that as palaeomagnetic data^{7–11} point to the Palaeozoic existence of Pangean and Pacific supercontinents at the sites of Africa–Atlantic and Pacific geoid highs, the latter are the still-subsiding relics of those produced by the Palaeozoic mantle's thermal expansion consequent on the protracted insulation provided by the supercontinents.

However, several reasons make this model unacceptable. The dimensions of Pacific and Africa–Atlantic geoid highs *vis-à-vis* those of the corresponding Pacific and Pangea supercontinents present one problem. The Pacific supercontinent comprised present east and south-east Asian and north-west North American blocks but reached hardly one-third the size of Pacific geoid high. This problem persists despite (1) adding the Siberian

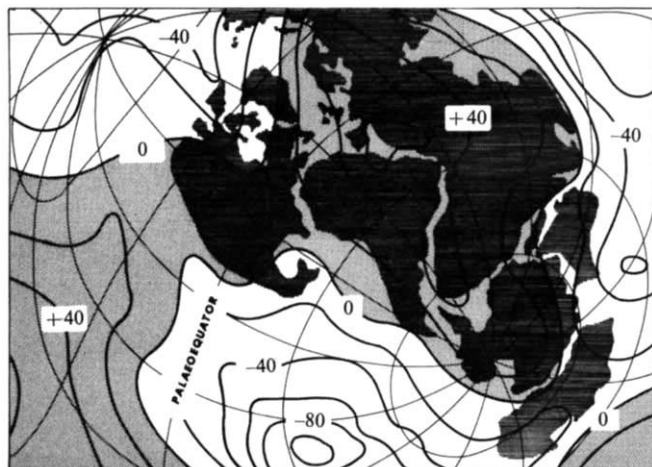


Fig. 1 The superposition of residual geoid anomaly map over that of Permo-Carboniferous Pangea. The position of Africa is common to the two maps. The geographical coordinates are for Permo-Carboniferous times. Geoid contours (in metres, geoid high region is shaded) correspond to the present position of Africa.

plate, which is believed¹²⁻¹⁵ to have joined European plate during the late Palaeozoic Uralian orogeny, to this assembly, (2) assuming larger initial sizes of east and south-east Asian blocks as they must have shrunk during the Mesozoic Indosinian and Yenshan orogenies, and (3) considering the presence of plateaus, ridges, rises and seamounts in the Pacific, most of which are demonstrably younger than this supercontinent's dispersal. Compared with the enormous geoid high produced by this small supercontinent, the much larger Pangea apparently had a weaker effect. Figure 1 illustrates this by superimposing Chase's map of residual geoid over Morel and Irving's¹¹ palaeomagnetic map of Permo-Carboniferous Pangea and shows Gondwanic Indo-Australian assembly and part of Antarctica as actually occupying the geoid low region. Using the hotspot reference frame, as Anderson has done, in place of the palaeomagnetic reference frame used here, offers little help. Thus, if continental insulation were to produce a lasting

geoid high then this instance of eastern Gondwanaland having an opposite effect is particularly intriguing when we recall that Gondwanic India-Australia-Africa assembly lasted throughout the late Proterozoic to late Mesozoic life of Gondwanaland^{16,17}. Further, the Pacific supercontinent was extinct by late Mesozoic times and only oceanic lithosphere now covers the Pacific geoid high whereas continental lithosphere still covers part of the Africa-Atlantic geoid high. This aggravates the problem as the presently comparable sizes of two anomalies suggest, in terms of Anderson's model, that the initial Pacific geoid high was much larger than the Africa-Atlantic geoid high. It would have been a different matter altogether if any of the present constituents of erstwhile Pacific supercontinent were to have abnormally highly radiogenic crust that could have augmented thermal expansion of the sub-Pacific mantle by adding to normal continental insulation. This is not the case either. These apart, the Antarctic and Indo-Siberian geoid lows also need to be explained. Even if the Antarctic and Indian plates are taken to have reached their present locations (the latter occupied the geoid low even in the Gondwanic times) too recently¹⁸ to have raised the geoid in their respective regions by way of continental insulation mechanism, Siberia has remained at about its present site ever since the late Palaeozoic times. Thus, although Anderson has not addressed this subject, the presence of an enormous geoidal depression in Siberia remains inexplicable in terms of his model.

These inadequacies in Anderson's continental insulation model apparently necessitate the search for an alternative mechanism to account for the locations of geoid highs at the sites of Palaeozoic supercontinents. However, this is not the only correspondence between palaeocontinental configurations and geoidal undulations. Figure 2 shows that a definitive correlation does indeed exist between the latter and the general pattern of middle and late Phanerozoic continental dispersals and accretions. Note that the Pacific geoid high is located where only divergent plate motions have occurred since the middle Palaeozoic, for example, the middle Palaeozoic rupture of Siberia from the Pacific supercontinent, its rapid westerly drift and the fusion with European plate, and the Mesozoic breakdown of this supercontinent and subsequent dispersals of its constituents to their present sites. Similarly, the Africa-Atlantic

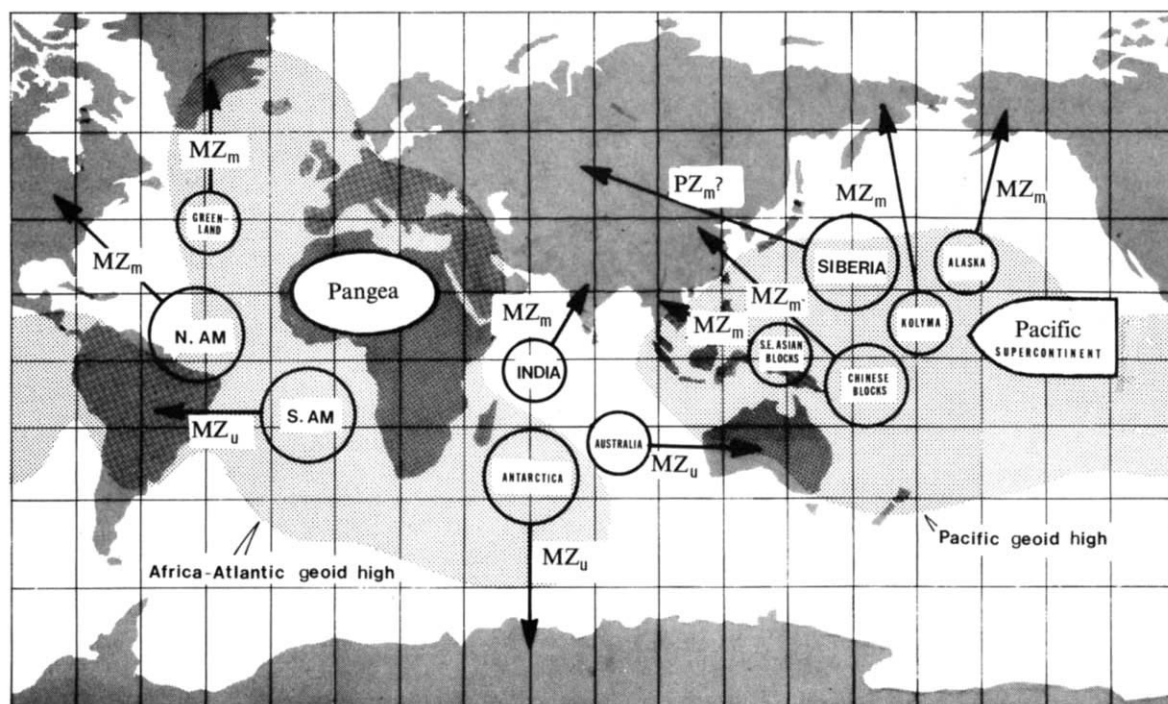


Fig. 2 A generalized schematic correlation of geoid anomalies (positive region is shaded, negative region is blank) and Palaeozoic continental configuration (circles denote palaeopositions, arrows denote shifts towards present positions, and geological dates: MZ_u, upper Mesozoic; MZ_m, middle Mesozoic; PZ_m, middle Palaeozoic, are the times of initial ruptures of the continental blocks from their supercontinental locations).

geoid high occupies the region that has witnessed divergent plate motions such as the middle Mesozoic drift of North America and Greenland from the rest of Pangea, late Mesozoic breakdown of the Gondwanic assembly and the drifts of its constituents to their present sites, and the early to middle Tertiary rotation of Eurasia relative to Africa. Compared with these divergent plate motions in geoid high regions, the geoid low region shows a conspicuously contrasting correlation with convergent plate tectonism. This is evidenced in the location of the Indo-Siberian low over the site where successive convergent motions have created today's Eurasian supercontinent¹⁵ and in that of the Antarctic geoid low over another palaeoconvergence zone. We also note, in this context, that present convergence displays an opposite pattern as narrow bands of geoid highs due to cold subducting slabs in these zones had to be removed⁴ to prepare the map of the residual geoid discussed here.

The geoid anomalies thus permit the separation of overall tectonism (for example, the correlation of global plate divergence and convergence regions with the rises and depressions of geoid) from the localized tectonism of plate boundaries (for example, geoid highs at consuming plate margins). Thermal convection in the mantle being the most accepted driving mechanism for plate motions^{19,20}, it seems appropriate to seek a convective origin for these anomalies. The observed duality of relation between regional and localized convergent tectonism and geoid anomalies then leads us to Chase's inference that the source of residual geoid anomalies probably lies in the lower mantle convection.

Figure 3 outlines a likely model for this. The geoidal rise is treated here as an upwelling caused by the upward pressure effect of lower mantle's ascending convection and the geoidal depression as a downwelling resulting from a suction effect of descending lower mantle convection. In the case of the former, the rising hot thermal plume at the centre of ascending convective shells causes the thermal expansion of upper mantle which in turn raises the geoid. Density contrast being negative at the lithosphere-asthenosphere contact^{21,22}, the resulting rise of hot asthenospheric material (1) aids the swelling of geoid, in isostatic compensation, by forming a low density 'root', (2) generates hotspots, and (3) sets in the divergent motions of rigid lithospheric blocks, ruptured at places of maximum deviatoric tensile strength, by creating new oceanic lithosphere along mid-ocean ridges. Lithosphere contributes to geoid's rise by providing thermal insulation, partly by its own lower thermal conductivity and conductive heat flow (heat flow in the underlying asthenosphere is convective) and partly by its higher intrinsic heat production particularly when the continental crust is present, to help the asthenosphere's expansion. The opposite process of descending lower mantle convection, as mentioned

before, then explains the geoidal depressions. Here, cold thermal plume at the confluence of descending convective shells causes downwelling of the lithosphere and thus accelerates convergent motions of lithospheric blocks by complementing their divergent motions in the adjacent regions of ascending convection. Note that this lithosphere is essentially oceanic compared with the continental lithosphere capping the regions of divergent tectonism at the onset of the process. The negative buoyancy of this lithosphere also helps the downwelling process and the result is one of depressing the geoid: hence the correlation between geoid lows and palaeoconvergent tectonism. As distinct from these two lower mantle processes which affect the entire overlying region, the higher harmonics of geoid can be explained as follows. The lithospheric fracture-filling mechanism above the ascending lower mantle convection, and the resulting undulations at the base of the lithosphere, explain (1) the presence of adjacent expansive zones with no intervening zone of subduction (such as the East African Rift, where an expansive zone²² separates the South Atlantic and Carlsberg and South-west Indian Ocean Ridges), and (2) the linear bands of geoid and gravity anomalies in the Pacific^{23,24} that are otherwise attributed to longitudinal convective rolls²⁵ in the sub-Pacific mantle. Similarly, part of the compaction of lithosphere above the region of descending lower mantle convection occurs by way of a simple fracture-slicing mechanism. This relatively shallow phenomenon causes localization of excess uncompensated mass of cold subducting slabs below folded mountain belts and deep sea trenches, and produces narrow bands of geoid highs (see, Figs 2 and 5 of ref. 4).

By thus relating long-wavelength geoidal undulations and long period global tectonism through the upper mantle repercussions of lower mantle convection, this rather simplistic model solves the problems encountered^{3,26,27} in locating a sub-crustal source for global gravity anomalies and incorporates the postulated thermal origin for mid-plate hotspots²⁸. More advanced models of the geoid²⁹ prove the reliability of geoid anomalies discussed here and the presence of lateral velocity variations in the lower mantle^{30,31} beneath Africa and Pacific lends added credence to this model. Although inconsistent with the models of upper mantle^{32,33} and whole mantle^{34,35} convection, our preference for lower mantle convection is supported by Olson's³⁶ study favouring deeper mantle convection terminating at a depth of 670 km. And yet, the subject of mantle convection being so speculative, the possibility of uncoupled two-layer convection^{4,37} or lithospheric recycling^{38,39} cannot be discarded in this context. In conclusion, these discussions suggest that a realistic appraisal of the material properties for mantle convection should be consistent not only with the observed surface fields⁴⁰ but also with the palaeogeodesic and palaeotectonic considerations enumerated here.

I thank the authorities at Montana College of Mineral Science and Technology, Butte, and the University of Ilorin for supporting my work and Professor Charles Wideman and Mr Richard Macmeekin for helpful discussions. I have also benefitted greatly from constructive criticisms by Dr Clement Chase.

Received 27 July 1982; accepted 24 March 1983.

- Briden, J. C., Hurley, A. M. & Smith, A. G. *J. geophys. Res.* **86**, 11631–11656 (1981).
- McElhinny, M. W. & Valencio, D. A. (eds) *Palaeoreconstructions of the Continents* (Vol. 2 Geodynamics Series, International Geodynamics Program, 1981).
- Anderson, D. L. *Nature* **297**, 391–393 (1982).
- Chase, C. G. *Nature* **282**, 464–468 (1979).
- Wagner, C. A., Lerch, F. J., Brown, J. E. & Richardson, J. A. *J. geophys. Res.* **82**, 901–914 (1977).
- Crough, S. & Jurdy, D. *Earth planet. Sci. Lett.* **48**, 15–22 (1980).
- McElhinny, M. W., Embleton, B. J., Ma, X. H. & Zhang, Z. K. *Nature* **293**, 212–216 (1981).
- Panushka, B. & Stone, D. *Nature* **293**, 561–563 (1981).
- Van der Voo, R., Jones, M., Gromme, C. S., Eberlein, G. D. & Churkin, M. *J. geophys. Res.* **85**, 5281–5296 (1980).
- Irving, E. *Nature* **270**, 304–309 (1977).
- Morel, P. & Irving, E. *J. geophys. Res.* **86**, 1858–1872 (1981).
- Dewey, J. F. & Bird, J. M. *J. geophys. Res.* **75**, 2625–2647 (1970).
- Hamilton, W. *Bull. geol. Soc. Am.* **81**, 2553–2576 (1970).
- McElhinny, M. W. *Palaeomagnetism and Plate Tectonics* (Cambridge University Press, 1973).
- Burrett, C. F. *Earth planet. Sci. Lett.* **21**, 181–189 (1974).

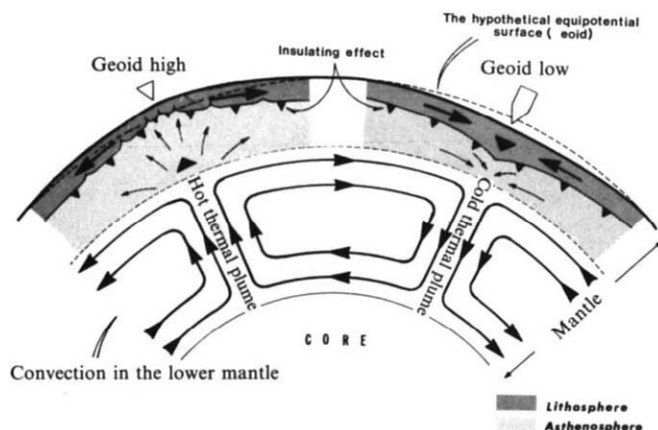


Fig. 3 Schematic outline of the model relating lower mantle convection, middle and late Phanerozoic plate motions, and the geoid anomalies described in the text.

16. Pal, P. C. *Curr. Sci.* **44**, 144–147 (1975).
17. Veevers, J. J. & McElhinny, M. W. *Earth Sci. Rev.* **12**, 139–159 (1976).
18. Embleton, B. J. J. & McElhinny, M. W. *Earth planet. Sci. Lett.* **58**, 141–150 (1982).
19. Oxburgh, E. R. & Turcotte, D. L. in *Rep. Prog. Geophys.* **41**, 1249–1312 (1978).
20. Morgan, W. J. in *The Sea* Vol. 7 (ed. Emiliani, C.) 443–488 (Wiley, New York, 1981).
21. Press, F. in *Gravity and Tectonics* (eds De Jong, K. A. & Scholten, R.) 7–16 (Wiley, New York, 1973).
22. Brown, C. & Girdler, R. W. *J. geophys. Res.* **85**, 6443–6455 (1980).
23. Marsh, B. D. & Marsh, J. G. *J. geophys. Res.* **81**, 5267–5280 (1976).
24. Marsh, J. G., Marsh, B. D., Williamson, R. G. & Wells, W. T. *J. geophys. Res.* **86**, 3979–3997 (1981).
25. Richter, F. M. & Parsons, B. J. *J. geophys. Res.* **80**, 2529–2541 (1975).
26. Higbie, J. W. & Stacey, F. D. *Nature phys. Sci.* **234**, 130–132 (1971).
27. Bott, M. H. P. *Earth planet. Sci. Lett.* **11**, 28–34 (1971).
28. Crough, S. T. *Geophys. J. R. astr. Soc.* **55**, 451–469 (1978).
29. Lerch, F. J., Klosko, S. M., Laubscher, R. E. & Wagner, C. A. *J. geophys. Res.* **84**, 3897–3916 (1979).
30. Dziewonski, A. M., Hager, B. H. & O'Connell, R. J. *J. geophys. Res.* **82**, 239–255 (1977).
31. Sengupta, M. K., Hassell, R. E. & Ward, R. W. *J. geophys. Res.* **86**, 3913–3934 (1981).
32. Richter, F. M. *J. geophys. Res.* **84**, 6783–6795 (1979).
33. McKenzie, D. & Jarvis, G. J. *J. geophys. Res.* **85**, 6093–6096 (1980).
34. Davies, G. F. *Geophys. J. R. astr. Soc.* **49**, 459–486 (1977).
35. O'Connell, R. J. *Tectonophysics* **38**, 119–136 (1977).
36. Olson, P. J. *J. geophys. Res.* **86**, 4881–4890 (1981).
37. McKenzie, D. & Weiss, N. *Geophys. J. R. astr. Soc.* **42**, 131–174 (1975).
38. Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421–436 (1982).
39. Chase, C. G. *Earth planet. Sci. Lett.* **52**, 277–284 (1981).
40. Kaula, W. M. *J. geophys. Res.* **85**, 7031–7044 (1980).

Magnetic variation anomalies in northern England and southern Scotland

R. J. Banks, D. Beamish* & M. J. Geake

Department of Environmental Sciences, University of Lancaster, Lancaster LA1 4YQ, UK

*Geomagnetism Unit, Institute of Geological Science, Murchison House, West Mains Road, Edinburgh EH9 3LA, UK

Single-station transfer functions linking the time variations of the vertical and horizontal components of the magnetic field at stations in northern England and southern Scotland have been compiled into a uniform data set. From hypothetical event analysis we show here that there are two distinct anomalies in the Borders region. One runs south-west to north-east, immediately to the south-east of the Southern Uplands Fault; the second follows the Northumberland Basin, and seems to exist because the conductive sedimentary rocks filling the basin create a link between the Irish and North Seas. If the Iapetus suture is marked by a conductivity anomaly, as has been suggested^{1,2}, these results place it beneath the Southern Uplands, unless it is masked by the surface conductor in the Northumberland Basin.

During the past 10 yr, several Geomagnetic Deep Sounding investigations of the electrical structure of the crust and upper mantle of southern Scotland and northern England have been described^{3–5}. The area which they cover in most detail is the Southern Uplands and the northern part of the Northumberland Basin. We have accumulated a large body of similar data covering northern England and the Borders—as yet unpublished⁶. Because of the importance which the area has acquired in reconstructions of the events associated with the closure of the Iapetus Ocean⁷, it seems desirable to reduce the data to a common form, so that a picture of the conductive structure of the entire region can be presented.

The only common form in which the data are available is as single-station transfer functions, linking the vertical and horizontal components of the magnetic field variations at specific frequencies, at each site⁸. Our results indicate that, in the period range 10–10⁴ s, transfer functions for the north of England are influenced by the magnetic fields associated with three different modes of induced current flow. At periods greater than 2,000 s, they depend on the direct fields of currents flowing in the Atlantic Ocean and in the shallow seas around the British Isles. Between 300 and 1,500 s, the fields at many sites are produced by the perturbation of currents driven

through the land as part of a thin-sheet induction process which includes the shallow seas. At periods <100 s, the transfer functions of sites which are sufficiently well insulated from conducting pathways between the seas detect currents produced by local induction in isolated regions of high conductivity. The largest data set that we could generate was at a period of 750 s and consisted of results from 94 stations; we anticipate that it will provide a picture of anomalous current concentrations in the shallow seas and, of more importance to us, in conductive 'links' in the land between them.

The technique that we have chosen to present the data is hypothetical event analysis⁹. In this approach, the single-station transfer functions are used to predict the vertical field of anomalous currents which would be produced when the horizontal component of the magnetic field across the region is spatially uniform, and has a specified polarization and intensity. Unfortunately, the prediction is wholly accurate only when the anomalous horizontal fields of the internal currents are negligible in comparison with the normal field. The presence of anomalous horizontal fields biases downwards the vertical fields computed using single-station transfer functions. In this case, we are forced either to accept the bias, or to discount any attempt to present a joint analysis of the data, as no other response measures are available for the area as a whole. However, at 35 stations on and around the Alston Block, inter-station transfer functions have been computed relative to a reference site at Durham (Fig. 1). The differences between the horizontal fields over this array and those at Durham, at a period of 750 s, rarely exceed 20% of the average value. Nonetheless, the knowledge that the bias exists should make us very cautious in interpreting the data, particularly in attempting to draw quantitative inferences from the hypothetical event maps.

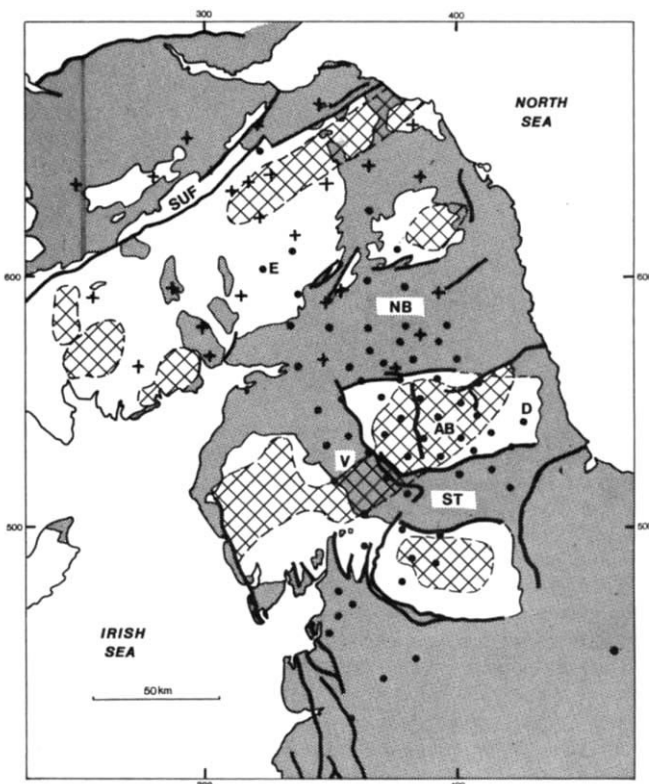


Fig. 1 Magnetometer sites used in the analysis: +, published data^{3,4,13}; ●, unpublished data; —, major faults. Cross-hatching indicates areas underlain by granite batholiths. Stippled areas are underlain by substantial thicknesses of post-Caledonian sedimentary rock. SUF, Southern Uplands Fault; NB, Northumberland Basin; AB, Alston Block; V, Vale of Eden; ST, Stainmore Trough; E, Eskdalemuir; D, Durham.

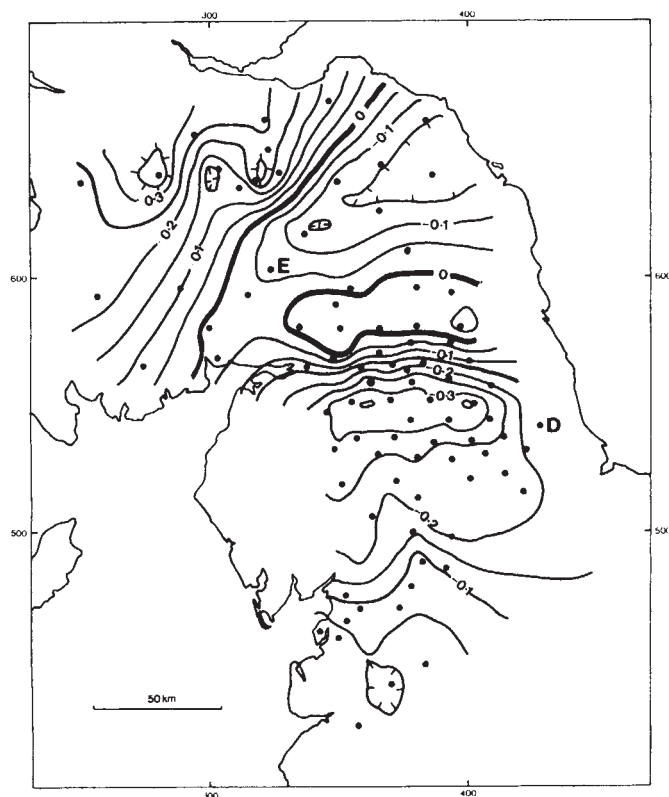


Fig. 2 Map of the in-phase part of the vertical magnetic field produced when the horizontal field is directed towards magnetic north (9° W), and has unit amplitude.

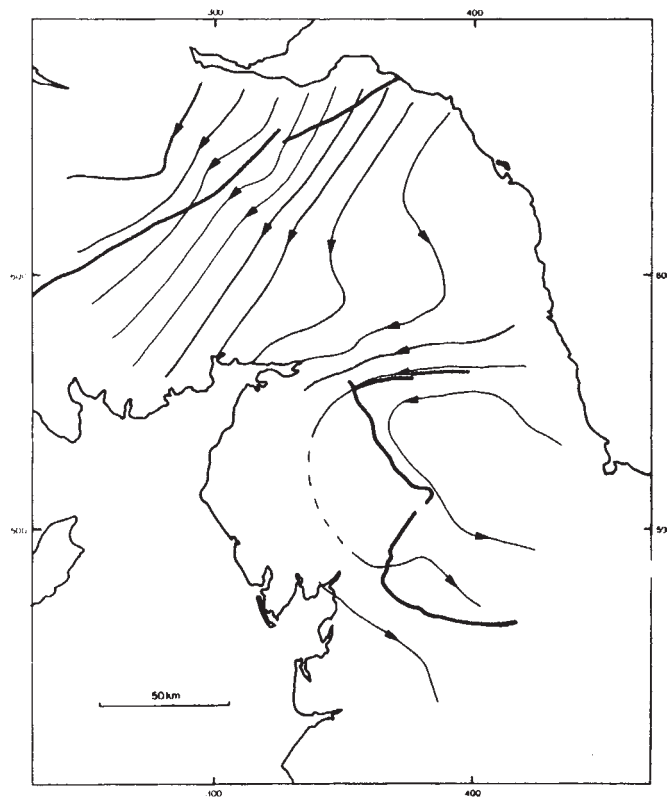


Fig. 3 The current stream function in a thin sheet at a depth of 5 km, derived from Fig. 2. Approximately 10 A flows between the stream lines when the amplitude of the horizontal magnetic field is 1 nT.

Figure 1 shows the distribution of sites in relation to the major structures in the crust which are likely to influence the flow of induced current. The granite batholiths should act as resistive blocks around which the current will be deflected, while it will be concentrated by the sedimentary basins.

The direction of the horizontal magnetic field which maximizes the anomalous vertical fields across the area as a whole is very close to magnetic north-south, and the anomalies are most strongly developed in the part of the vertical field which is in-phase with the horizontal. Figure 2 shows a contour map of the in-phase vertical field associated with a northward horizontal field of unit amplitude. The expected direction of the regional current flow for this azimuth of the magnetic field is from east to west. Note that a uniform regional current of this kind, whether it is restricted vertically or distributed with depth, produces no vertical field; the observed anomalous field is entirely due to the perturbations in the flow which are created by horizontal contrasts in conductivity. If we assume that the anomalous currents are concentrated in a thin sheet at a specific depth, the current stream function can be calculated directly from the map of anomalous vertical fields¹⁰. Figure 3 shows the equivalent current system at a depth of 5 km derived from the vertical field map in Fig. 2. Because of the limitations in the analysis (the effect of anomalous horizontal fields that we have been forced to ignore, and the fact that the current is not necessarily concentrated at a single depth), Fig. 3 must be regarded as an aid to visualization of the anomalous structures, rather than a basis for quantitative interpretation.

The most interesting features in Fig. 3 are those regions where the sheet of current is concentrated by narrow belts of relatively highly conducting rock in the crust. It appears that this polarization of the source field does not give rise to detectable fields arising directly from current concentrations in the shallow seas. Instead, the current is enhanced in two principal linear features through the land: one which runs from north-east to south-west across the Southern Uplands, with its axis to the

south-east of the Southern Uplands Fault, and the second which runs from east to west through the Northumberland Basin, between the North Sea and the Solway Firth.

Previous workers³⁻⁵ have defined and commented on the first of the two anomalies. Its direction is parallel to that of the Caledonian structures, but it does not coincide with any known belt of conductive rock close to the surface. Unfortunately, the station spacing in this region is often quite large and irregular, and the vertical field map does not impose very strong constraints on the depth of the current system. Magnetotelluric measurements have been made in the Southern Uplands¹¹, but, because they cover a relatively narrow band of frequencies, their resolution of upper crustal structure is poor. In the lower crust and upper mantle, they indicate a conductive zone at depths between 20 and 70 km. The cause of the high conductivity is unclear, although it has been suggested that it may be due in part to the presence of hydrated rocks and dehydration at the amphibolite-granulite transition^{12,13}.

Although the Northumberland Basin anomaly merges with the Southern Uplands anomaly at its western end, it is completely separate in the east. Note that the magnetic observatory at Eskdalemuir, where the anomalous character of the variation fields was first detected, lies where the anomalous fields of the two structures merge. The Northumberland Basin anomaly is very well defined by the dense network of IGS/Lancaster stations. The southern edge of the anomalous current concentration coincides with the northern margin of the Weardale granite, which underlies the Alston Block. When the current sheet is placed at depths in excess of 5 km, the current stream function becomes unstable, indicating that at least part of the actual current is shallower than 5 km. Magnetotelluric measurements at a site in the Northumberland Basin¹¹ also indicate a shallow conductor, although they are not conclusive because the sounding is close to the southern margin. The variation in depth of the pre-Carboniferous basement within the basin has been established by seismic refraction and gravity surveys¹⁴.

The sediments are at their thickest (up to 3 km) towards the southern margin, where the current density is greatest. Taken together, these facts suggest that the current responsible for the magnetic variation anomaly is concentrated principally in the relatively conductive Carboniferous sedimentary rocks of the Northumberland Basin.

To the west of the Alston Block, in the Vale of Eden, is a similar sedimentary basin which runs in a north-south direction, and which contains up to 3 km of Carboniferous and Permo-Triassic sediment¹⁵. If these sedimentary basins responded to electromagnetic fields as isolated conductors, the maximum response of the Vale of Eden should occur when the regional current flow is north-south, that is the horizontal magnetic field is east-west. However, at 750 s period, the basin responds to neither azimuth of the field, in striking contrast to the Northumberland Basin. What differentiates the two basins are the links which the Northumberland Basin makes at either end to the laterally extensive conductors of the Irish and North seas, together with their own associated sedimentary basins. The Vale of Eden is isolated at its southern end, either by the postulated ridge connecting the Lake District and Weardale

granites¹⁵, or by the westward thinning of sediments in the Stainmore Trough.

It has been suggested that the Northumberland Basin marks the Iapetus suture, the line of closure of the proto-Atlantic Ocean⁷. Before it can be established that there is a deep conductive structure beneath the basin, it will be necessary to model the effects of the currents in the shallow sedimentary rocks, using the available data on their thickness. We believe that, if the response of the Basin to electromagnetic fields with periods of 750 s is to be correctly modelled, it will be essential to include the Irish and North Seas and their underlying sediments. Without the results from such modelling, all we can say at present is that a deep conductor beneath the basin does not seem essential. The only completely unexplained feature of the maps is the Southern Uplands anomaly, and, if we believe that the line of the suture must be marked by a conductivity anomaly, this must be reckoned the most likely candidate.

We gratefully acknowledge the support of the NERC in the form of grant GR3/2371 and a grant from the Deep Geology Project. This paper is published with the approval of the Director, Institute of Geological Sciences (NERC).

Received 7 February; accepted 5 April 1983.

1. Law, L. K. & Riddiough, R. P. *Can. J. Earth Sci.* **8**, 1094-1106 (1971).
2. Bailey, R. C. & Edwards, R. N. *Geophys. J. R. astr. Soc.* **45**, 97-104 (1976).
3. Hutton, V. R. S. & Jones, A. G. *Adv. Earth planet. Sci.* **9**, 141-150 (1980).
4. Hutton, V. R. S. *et al. Phys. Earth planet. Inter.* **24**, 66-87 (1981).
5. Edwards, R. N., Law, L. K. & White, A. *Phil. Trans. R. Soc.* **270**, 289-323 (1971).
6. Grimes, D. I. F. thesis, Univ. Lancaster (1977).
7. Anderton, R., Bridges, P. H., Leeder, M. R. & Sellwood, B. W. *A Dynamic Stratigraphy of the British Isles* (Allen & Unwin, Edinburgh, 1979).

8. Banks, R. J. *Phys. Earth planet. Inter.* **7**, 339-348 (1973).
9. Bailey, R. C., Edwards, R. N., Garland, G. D., Kurtz, R. & Pitcher, D. J. *Geomagn. Geoelect.* **26**, 125-146 (1974).
10. Banks, R. J. *Geophys. J. R. astr. Soc.* **56**, 139-157 (1979).
11. Jones, A. G. & Hutton, V. R. S. *Geophys. J. R. astr. Soc.* **56**, 329-368 (1979).
12. Hutton, V. R. S., Ingham, M. R. & Mibipom, M. W. *Nature* **287**, 30-33 (1980).
13. Ingham, M. R. & Hutton, V. R. S. *Geophys. J. R. astr. Soc.* **69**, 595-606 (1982).
14. Lee, M. K. *Geophys. J. R. astr. Soc.* **69**, 281 (1982).
15. Bott, M. H. P. *J. geol. Soc. Lond.* **113**, 93-117 (1974).

The first fossil ctenophore from the Lower Devonian of West Germany

George D. Stanley Jr

Department of Geology, University of Montana, Missoula, Montana 59812, USA

Wilhelm Stürmer

Burgbergstrasse 20, D-8520 Erlangen, FRG

Of the 21 or so generally accepted animal phyla, only one, the Ctenophora, has no fossil record¹. The Phylum Ctenophora or comb jellies, are typically globular, swimming marine organisms, widely distributed in today's oceans. They are small, delicate, mostly planktonic animals of a gelatinous nature² that were first considered to be coelenterates but later were recognized as a distinct group of phylum rank. We have now discovered the pyritized fossil remains of a small, globular, soft-bodied ctenophore. It comes from the Lower Devonian Hunsrück Slate of West Germany and provides the first tangible evidence for the existence of the group. Our discovery demonstrates that the basic ctenophore body plan has changed very little over the past 400 Myr and suggests that the origin of the phylum must extend even further back in time.

Although poorly preserved, the shape, symmetry and anatomical details of the fossil described below leave little doubt that it is a ctenophore. The specimen was discovered during radiographic study of a variety of fossils in the collection of Günther Brassel found in Hunsrück Slate from the Eschenbach Slate Mine, Bundenbach, FRG. Millions of fossils from this formation include over 370 different invertebrate taxa³. The use of X rays in the study of other Hunsrück fossils has greatly expanded the resolving power of palaeontological observations, revealing details as fine as 5 µm (refs 4, 5). The preservation of soft tissues frequently occurs through the conversion of organic sulphur compounds to iron sulphides^{5,6}. Jellyfish-like hydrozoans (Velellidae) with exquisite details of the tentacles,

organs and soft tissues are preserved in a similar manner^{7,8}. The discovery of a ctenophore in the Hunsrück Slate establishes the existence of this phylum over 400 Myr ago in the early Devonian.

Systematic description: *Paleoctenophora brasseli* Stanley and Stürmer, n. gen., n. sp. (Figs 1 and 2). The description of the specimen is based on stereoradiographs. It is biradially symmetrical, oval-shaped, 13 mm high and 9 mm in diameter. The wall is exceptionally thin and discontinuous and there is a slightly depressed apical (aboral) region, strongly differentiated into distinct spherical granules (0.20-0.54 mm diameter) from which extend at least seven or possibly eight radially disposed linear features. The longest extends three-quarters of the body length. Others are short and thicken away from the apex, gradually dissipating and losing integrity. Fibrous, irregular matter is distributed within the specimen. Opposite the apex, the wall of the oral region has slight but perceptible indentation. On opposite sides, two tentacular projections emerge through the wall. One is short and thick with possibly an internal sac, the other is longer, extending 50 mm outside the body, with a few short branches along its length.

The presence of tentacles and apparent tentacle sheaths place this ctenophore in the Class Tentaculata; the shape, symmetry and details of the tentacles allow its assignment to the Order Cydippida. Finer details are insufficient for assignment to a particular family but the resemblance to *Pleurobrachia* is striking. The specimen is in the Bavarian State Collection, Munich, FRG (BSP 1938 ROEM 134). Original radiographs (WS 1021) are with W.S.

Only two doubtful fossil ctenophores are known⁹. These are *Rangea*¹⁰ and *Xenusion*¹¹. The former is now known to be a soft-bodied coelenterate of the Precambrian Ediacara fauna belonging to the Pennatulacea^{12,13}. The latter is undoubtedly a segmented Cambrian arthropod. The eight row combs of cydippids are very delicate external structures that are lightly attached to the gelatinous body. Each is composed of overlapping plates of fused cilia and their coordinated beating is the principal means of locomotion. Because of their delicate nature, it is not surprising to find the comb structures of *Paleoctenophora* in disarray. Relative to the body, they were more rigid,

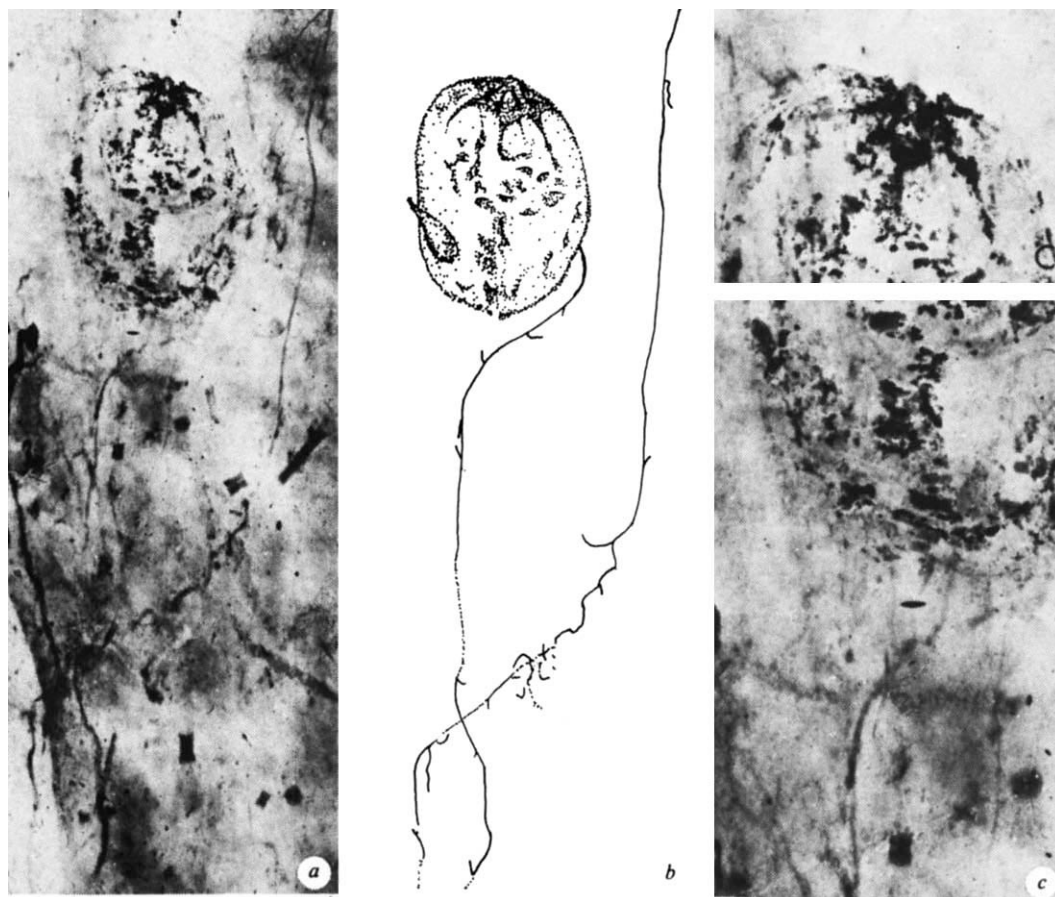
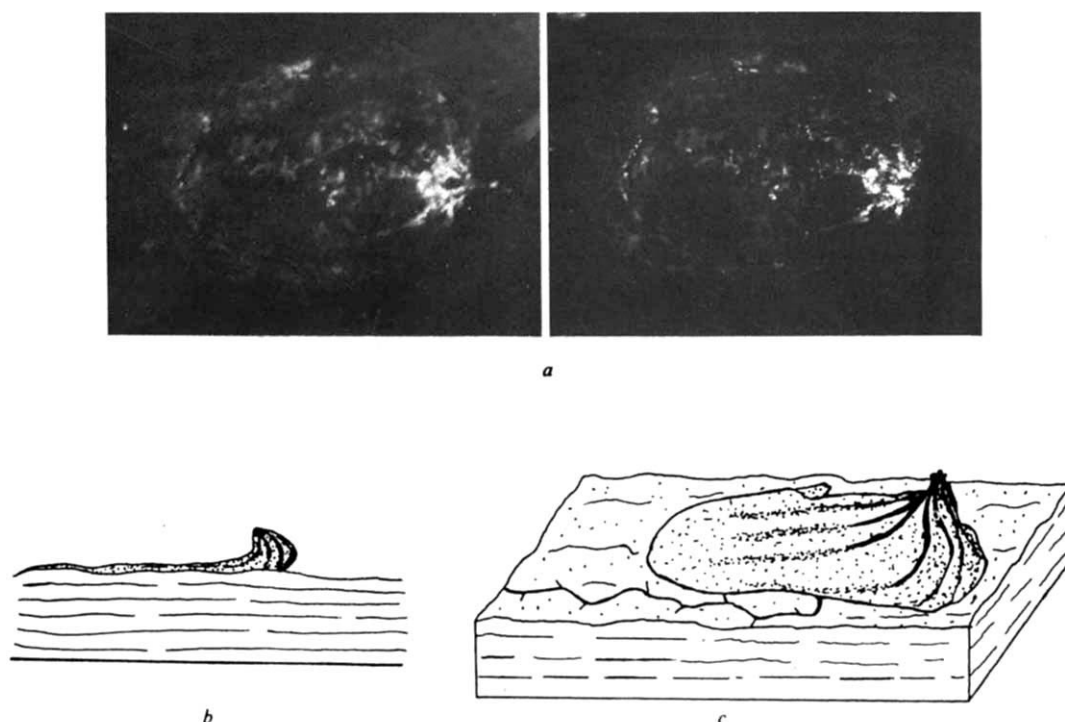


Fig. 1 *a*, *Paleoctenophora brasseli* n. gen., n. sp. Radiograph of fossil ctenophore (WS 1021) from the Hunsrück Slate of West Germany showing general shape of specimen and the associated structures lying in the matrix. A fine sinuous feature extending from the specimen is interpreted as the remains of a tentacle. A few short filaments branch off at various points along its length. To right of specimen, a similar structure is unattached but could also be the remains of a tentacle ($\times 2.3$). *b*, Sketch of specimen at left clarifying principal features. *c*, Top, enlarged details of apical region showing concentration of pyritic matter interpreted as remains of the statocyst and the radially extending comb rows ($\times 4.7$). *d*, Below, enlarged details of lower left part of specimen showing the protruding tentacle that emerges through the body wall. A trace of a sac-like structure, probably the remains of an internal tentacle sheath, is visible (upper left). Part of the right tentacle, showing indications of fine-branched extensions, is also present.

Fig. 2 *a*, Radiograph stereo-pair of *Paleoctenophora* showing flattened specimen entombed within the slate matrix ($\times 3.0$). Angle of stereo photographs accentuates slight three-dimensional details and emphasizes the relief in the apical region of the specimen. Original rigidity of comb rows may have reduced flattening in this area. *b*, Schematic drawing with proportions exaggerated showing a profile of the specimen and the raised apical region where comb row structures are present. *c*, Generalized sketch of the specimen based on stereo radiography giving orientation in the matrix.



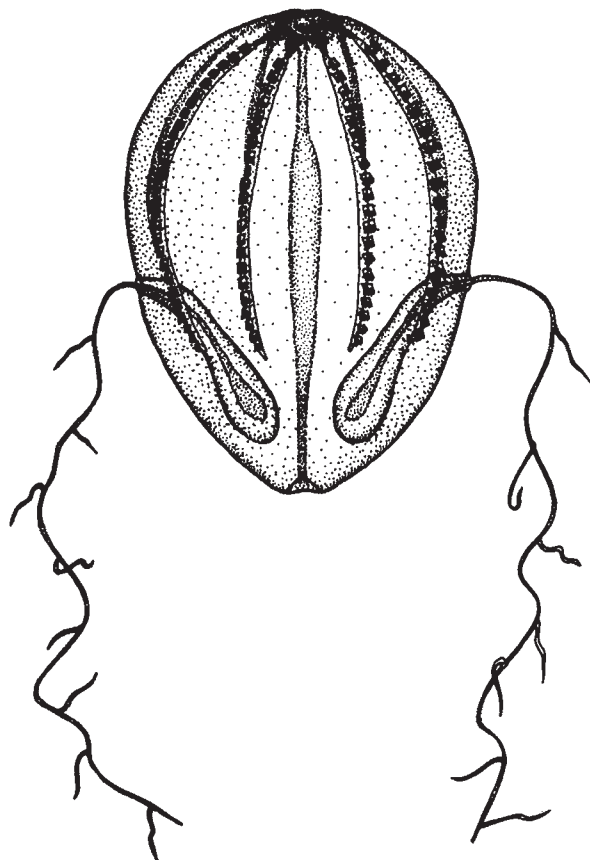


Fig. 3 Life reconstruction of *Paleoctenophora* showing principal anatomical features suggested by the fossil. Drawing illustrates the apical statocyst from which radiate eight longitudinal comb rows. Also shown are the paired, retractable tentacles with internal tentacle sheaths. In the centre, a part of the gastrovascular system runs through the body and terminates in the mouth. Many details of the internal anatomy, including much of the gastrovascular system, are not visible owing to lack of preservation in the fossil. Drawing illustrates the striking similarity with living cydippid ctenophores such as *Pleurobranchia*.

supporting the soft tissue (Fig. 2). The comb rows terminate in and attach to the aboral region (Fig. 2). Comparison with modern cydippids shows this to be the region of the statocyst or balance organ that in living ctenophores consists of minute calcareous spherules (statolith)¹⁴. Remains of such spherules are present in the specimen (Fig. 1c) and the opposite (oral) side reveals an indentation indicating a mouth (Fig. 1a). The fine, dark, pyritic matter scattered within the specimen (Fig. 1) could be degraded remains of muscle fibres, the pharynx and gastrovascular system.

Of the two tentacles indicated, one on the left side appears to be either broken off or retracted and is enclosed in a sac-like tentacle sheath (Fig. 1c). On the opposite side a tentacle emerges from the body and can be traced along a sinuous path for about 5 cm (Fig. 1a). Along its length, it displays several branched filaments. Other sinuous lines not connected with the specimen lie in different planes and are interpreted as worm burrows. Figure 3 illustrates a reconstruction of *Paleoctenophora*.

With only one fossil specimen representing an entire living phylum, *Paleoctenophora*, must be the rarest of all fossils. Very special conditions of burial account for the preservation of such a delicate, jelly-like organism. We have no idea how common ctenophores were in Devonian seas but today they are exceedingly abundant and diverse plankton feeders, forming an important component of oceanic marine ecosystems¹⁵.

The basic cydippid body plan of *Paleoctenophora* reinforces earlier evolutionary hypotheses based on embryological grounds. Since all ctenophores, including the most morphologi-

cally diverse, go through an early free-swimming cydippid larval stage, the Order Cydippida is postulated to be the ancestral form^{1,14}. *Paleoctenophora* therefore adds credence to this idea.

Although caution is necessary in a radiographic study, we judge the obvious similarities to living cydippid ctenophores convincing enough to justify our interpretations. The striking similarity of *Paleoctenophora* with living counterparts such as *Pleurobranchia* indicates the antiquity of the basic ctenophore body plan¹⁶. The presence of this grade of metazoan organization over 400 Myr ago, suggests that if ctenophores had a common origin with Cnidaria (Coelenterata) or any metazoans of the Radiata branch, it must extend even further back into the Palaeozoic, perhaps to the dramatic appearance of metazoans in the late Precambrian¹⁷.

We thank Professor R. Siewing, University Erlangen for his cooperation and for obtaining living specimens for our study, and Mr G. Brassel for donating this unusual specimen. G.D.S. acknowledges a US/West German Senior Fulbright Research Grant and W.S., DFG and Volkswagen Foundation. Drs Paul Cornelius, Brian Rosen and Richard Jeffries offered comments on our manuscript.

Received 21 December 1982; accepted 16 March 1983.

1. Geczy, B. *Acta biol. Acad. Sci. Hung.* **31**, 461–464 (1980).
2. Bayer, F. M. & Owe, H. B. *The Free-Living Lower Invertebrates* (Macmillan, London, 1968).
3. Mittmeyer, H.-G. *Natur Mus., Frankf.* **110**, 34–39 (1980).
4. Stürmer, W. *Natur Mus., Frankf.* **110**, 3–18 (1980).
5. Stürmer, W. & Bergström, J. *Paläont. Z.* **47**, 104–141 (1973).
6. Stürmer, W. *Science* **170**, 1300–1302 (1970).
7. Rauff, H. *Paläont. Z.* **21**, 194–213 (1939).
8. Yochelson, E. L., Stürmer, W. & Stanley, G. D. Jr *Paläont. Z.* (in the press).
9. Wells, J. W. & Hill, D. in *Treatise on Invertebrate Paleontology Pt F* (ed. Moore, R. C.) F478 (Geological Society of America and University of Kansas Press, 1956).
10. Gürich, G. *Int. geol. Congr. S. Af.* **1929** **15**, 670–680 (1930).
11. Moore, R. C. in *Treatise on Invertebrate Paleontology Pt O* (ed. Moore, R. C.) 016–021 (Geological Society of America and University of Kansas Press, 1959).
12. Glaessner, M. F. *Nature* **183**, 1472–1473 (1959).
13. Glaessner, M. F. *Scient. Am.* **204**, 72–78 (March) (1961).
14. Hyman, L. H. *Invertebrates: Protozoa through Ctenophora* (McGraw-Hill, New York, 1940).
15. Harbison, G. R., Madin, L. P. & Swanbert, N. R. *Deep-Sea Res.* **25**, 233–255 (1978).
16. Barrington, E. J. W. *Invertebrate Structure and Function* (Nelson, Sunbury-on-Thames, 1979).
17. Scrutton, C. T. in *The Origin of Major Invertebrate Groups* (ed. House, M. R.) 161–207 (Academic, New York, 1979).

Controlling the duration of photosynthetic charge separation with microwave radiation

Michael R. Wasielewski, Christian H. Bock, Michael K. Bowman & James R. Norris

Chemistry Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

Electron transfer reactions are commonplace in both chemical and biochemical transformations. These reactions generally involve the formation of transient radical ion pairs. Non-adiabatic electron transfer results in formation of a correlated radical ion pair¹. Only when the radical ions have separated to distances of the order of 10 Å are the internal magnetic fields in the radical ions able to perturb this correlation. Thus radical pairs initially born in a singlet state will develop triplet character as a function of time, and vice versa. Because the ensuing chemistry of the radicals is dependent on their spin state, external perturbations of the radical pair spin state may permit some control over the course of the chemistry. Microwave radiation can provide a strong internal magnetic field perturbation on a radical pair such as the one that results from photoinduced charge separation in bacterial reaction centres. We show here how this perturbation can be used to control the lifetime of the initially formed radical pair.

Photoexcitation of the reaction centre protein from purple photosynthetic bacteria results in rapid (<5 ps) formation of a radical pair, P[•] composed of an oxidized bacteriochlorophyll a dimer, P⁺ and a reduced bacteriopheophytin a molecule, I[•]

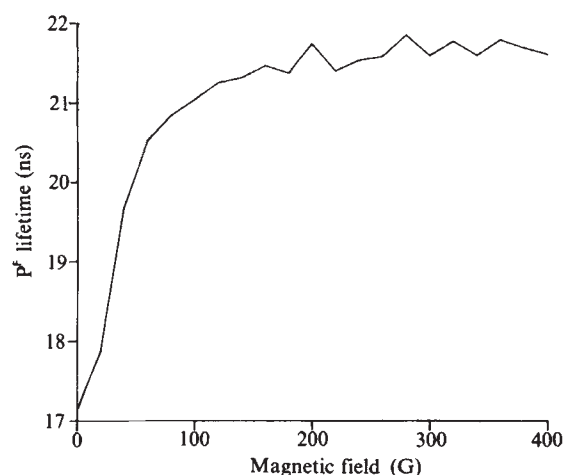


Fig. 1 P^F lifetime plotted against magnetic field for reaction centres from R-26 *R. sphaeroides*. Low magnetic field with no microwave radiation.

(ref. 2). If the endogenous quinone molecules in the protein are either removed or chemically reduced before excitation, P^F lives for about 20 ns in the absence of an external magnetic field³. During this time the initially formed singlet population of P^F , $^1[P^+I^-]$ is under the influence of local magnetic fields primarily due to nuclear hyperfine interactions in P^+ and I^- . As a result a fraction of the $^1[P^+I^-]$ population undergoes radical pair intersystem crossing (RP-ISC) to yield $^3[P^+I^-]$. Thus, the total population of P^F becomes a mixture of $^1[P^+I^-]$ and $^3[P^+I^-]$ the composition of which varies with time.

$^1[P^+I^-]$ decays directly to the singlet ground state with a rate constant k_s , while $^3[P^+I^-]$ back reacts to form 3P with a rate constant k_r . 3P lives for microseconds before intersystem crossing back to ground state singlet P . At zero magnetic field the $^1[P^+I^-]$ and $^3[P^+I^-]$ states mix because they are nearly degenerate. On application of a magnetic field one of the three triplet sublevels of $^3[P^+I^-]$, T_0 remains nearly degenerate with $^1[P^+I^-]$, while the remaining two levels T_{+1} and T_{-1} increase and decrease in energy, respectively⁴. In general, both in the presence and absence of an external magnetic field the mixed radical pair states are at least slightly paramagnetic during most of the lifetime of P^F . If during the 20-ns lifetime of the radical pair we can perturb the system by applying microwave radiation that possesses energy sufficient to induce transitions between the mixed radical pair states, we should be able to change the chemistry of the radical pair between that due to the formation of triplet products and that due to the formation of singlet products. Moreover, if the respective rate constants, k_i and k_s , of these radical pair decay processes are different we should also be able to control the lifetime of the radical pair by application of microwave radiation.

Reaction centres were isolated from the R-26 mutant of *Rhodospseudomonas sphaeroides* and depleted of endogenous quinones³. A 80–100 μ M solution of reaction centres was placed in a standard electron paramagnetic resonance (EPR) flat cell in a Varian optical transmission cavity centred between the poles of an electromagnet. One microsecond microwave pulses possessing powers up to 20 kW were provided by a magnetron source. The optical absorbance of P^F was monitored with 420 nm light from a flashlamp during the time following a 0.5–1.0 mJ, 6 ns, 600-nm dye laser pulse. Optical density changes were monitored as a function of time with a 2.5 ns response time photomultiplier. Typically, 256 laser shots were averaged at each field value. Each field step was 5 G. The entire optical observation was initiated and completed within the 1 μ s microwave pulse period, so that the sample was under constant microwave irradiation during this time. Magnetic field sweep and data acquisition were under computer control. The decay of the optical transient at each field value was analysed using the unrestricted Fourier transform technique of Provencher⁵.

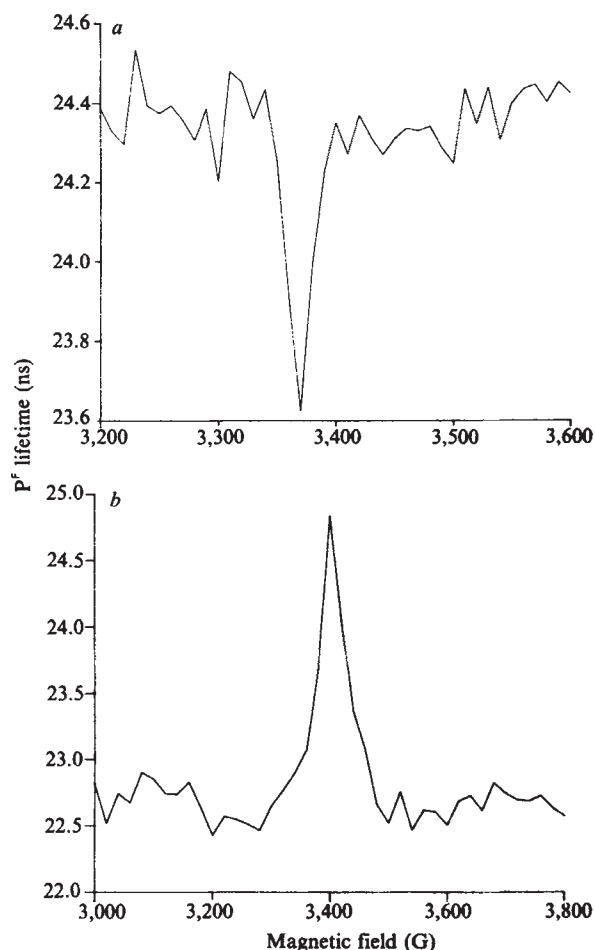


Fig. 2 *a*, Lifetime detected magnetic resonance spectrum of P^F in reaction centres from R-26 *R. sphaeroides*. 1-kW microwave radiation at 9.4 GHz applied. *b*, Lifetime detected magnetic resonance spectrum of P^F in reaction centres from R-26 *R. sphaeroides*. 7-kW microwave radiation at 9.4 GHz applied.

Figure 1 shows a plot of P^F lifetime as a function of magnetic field out of resonance. The lifetime of P^F increases with increasing magnetic field by ~25% (similar behaviour has been reported previously³). Application of a magnetic field splits the three triplet sublevels of $^3[P^+I^-]$ leaving only T_0 nearly degenerate with $^1[P^+I^-]$. As a consequence RP-ISC to T_{+1} and to T_{-1} is greatly slowed resulting in the lifetime of P^F being controlled primarily by decay to ground state.

Two states possessing only partial triplet character result from mixing $^1[P^+I^-]$ and T_0 within P^F . To observe a P^F resonance microwave transitions must occur between these mixed states and T_{+1} and T_{-1} . This requires that a substantial H_1 field be present during the lifetime of P^F .

In Fig. 2*a* the P^F lifetime exhibits resonance at a magnetic field where $g = 2$ when microwaves are present. At moderate powers, 1 kW, the lifetime of P^F decreases at resonance. Resonant microwaves increase the amount of $^3[P^+I^-]$ at the expense of that of $^1[P^+I^-]$. The resultant decrease in P^F lifetime shows that $k_i > k_s$. If $k_i = k_s$, then no resonance would be observed. Finally, if $k_s > k_i$, turning on RP-ISC with microwave radiation would increase the lifetime of P^F .

If the microwave power is increased so that H_1 is larger than the field due to the hyperfine interaction that mixes $^1[P^+I^-]$ and $^3[P^+I^-]$, then the spins will remain locked in the phase relationship to each other that they possessed when the radical pair was born⁶. Thus, at large values of H_1 in resonance we turn off RP-ISC. This results in a larger fraction of the radical pair population remaining in the singlet state. The observed increase in the lifetime of P^F reflects the dominance of k_s in the decay of P^F in these conditions and confirms the fact that $k_i > k_s$.

Our results show that the duration of photosynthetic charge separation can be controlled with microwave radiation. Moreover, the exercise of this control makes it possible to observe the dynamics of the radical pair processes that comprise the primary photochemistry of bacterial photosynthesis.

We thank David Budil for the preparation of the R-26 reaction centres. This work was supported by the Division of Chemical Sciences, Office of Basic Energy Sciences, US Department of Energy under contract W-31-109-Eng-38.

Received 7 February; accepted 25 March 1983.

1. Closs, G. L. *Adv. magn. Reson.* **7**, 157–229 (1974).
2. Holten, D. *et al. Biochim. biophys. Acta* **592**, 461–477 (1980).
3. Schenck, C. C., Blankenship, R. E. & Parson, W. W. *Biochim. biophys. Acta* **680**, 44–59 (1982).
4. Thurnauer, M. C., Katz, J. J. & Norris, J. R. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3270–3274 (1975).
5. Provencher, S. W. *J. chem. Phys.* **64**, 2772–2777 (1976).
6. Slichter, C. P. *Principles of Magnetic Resonance*, 214 (Springer, Berlin, 1978).

ABO genes are differentially distributed in socio-economic groups in England

J. A. Beardmore & F. Karimi-Booshehri

Department of Genetics, University College of Swansea, Singleton Park, Swansea SA2 8PP, UK

No direct evidence is available concerning what average genetic differences, if any, characterize the segments of socially stratified human populations, although theoretical considerations^{1–4} suggest that genetic differentiation within such populations is to be expected. We have now analysed two large samples of data from blood donors in England to test whether the distributions of the ABO and Rhesus blood group phenotypes are random with respect to socio-economic groups as determined by occupational classification. We have found that in both native and migrant sections of the populations of two widely separated regions (south-west England and part of Yorkshire) and in both sexes, the A phenotype is highly significantly more, and the O phenotype significantly less, frequent than expected in social classes I and II, while the converse is seen in social classes III–V. An individual of the A phenotype has thus about 15% greater probability than chance would dictate of being placed in classes I and II. The distribution of the Rh⁺ and Rh[−] phenotypes does not differ significantly between classes. It seems unlikely that this nonrandom distribution of the ABO phenotypes among socio-economic groups results from sampling, historical or migrational effects and we conclude that the observed association is likely to result from pleiotropic effects of the ABO alleles (or genes closely linked to them) on attributes influencing occupational type, social mobility and social class.

Considerable data on ABO blood group phenotypes and alleles now exist showing marked differentiation between geographical areas such as Europe and Asia⁵, clines within areas such as the United Kingdom⁶ and some microgeographical heterogeneity related to invasion, settlement and migration patterns⁷. Associations of different ABO phenotypes with a variety of diseases have also been described⁸ and some authors⁹ believe differential resistance to diseases of historical significance to be a major cause of the existence of this polymorphism. However, apart from two studies^{10,11}, little effort has been made to analyse defined populations for heterogeneity of blood group distribution such as could arise from association of particular phenotypes with specific social, economic or cultural groups.

The work described here was made possible by the existence of a very large body of data derived from the UK Blood Transfusion Service Records for the period 1953–63, assembled by Dr A. E. Mourant and kindly made available to us by Dr D. Tills. In these data, Rhesus and ABO blood types, age, sex, birthplace, place of sample collection and occupation were

recorded for approximately 400,000 individual blood donors drawn from all parts of England. Two large samples from Yorkshire (mainly West Riding) and south-west England (Devon and Cornwall) totalling almost 14,000 have been analysed in detail to see whether the blood group phenotypes are randomly distributed with respect to occupational classes and in particular, between more skilled/creative versus less skilled/repetitive occupations.

These areas exclude samples from large cities. Our data for south-west England include all those previously used by Kopeč⁶, while for Yorkshire we have used all of the samples from the following places: Halifax, Huddersfield, Ilkley, Otley, Pickering, Ripon, Settle, Shipley and Skipton. Natives are defined as those individuals whose birthplace and place of residence (town or village) at the time of sampling were the same, all other individuals being classed as migrants. Many migrants would thus have moved only a short distance. Classification of individuals as to socio-economic group followed broadly the Registrar General's scheme¹². However, a restrictive view was taken as to the boundary between classes I and II such that where any doubt existed as to the appropriate class for any individual, he was scored as class II. Furthermore, to increase accuracy of classification, individuals with no, or poorly defined, occupational classification, such as engineer, civil servant, student or housewife, were not included in the analysis. The elimination of such individuals resulted in the loss of some 4,300 from the total sample but this does not significantly change the phenotypic distributions previously reported for these areas. Kopeč's values⁶ for the frequency of the O phenotype are 44.48 and 45.02% for the same two regions of south-west England and Yorkshire, while our figures (after exclusion) are 44.44 and 45.34% respectively.

The right-hand column of Table 1 shows the ABO phenotype distribution for natives and migrants for the two geographical areas. The great majority of migrants were derived from the United Kingdom, the remainder coming from the Republic of Ireland (about 1%) and (in a few cases) other European countries. The proportion derived from other ethnic groups (judged by surname and/or birthplace) is very small (no more than 1%) and such individuals have not been excluded from the analysis. The two areas differ in the frequency of blood groups (notably B), but χ^2 analysis shows that, within each area, there is no heterogeneity of overall blood group frequencies between the two sets of natives and migrants (south-west England $\chi^2_3 = 0.21$, $P > 0.97$; Yorkshire $\chi^2_3 = 0.48$, $P > 0.5$).

Table 1 gives the frequency distribution of ABO phenotypes by socio-economic class for the four sets of donors. The blood group frequencies in the right-hand column of Table 1 provide a basis for calculating the expected number of each of the ABO phenotypes in each socio-economic group and the deviations of the observed from expected numbers. The signs of these deviations are also given in Table 1. It is clear that there is a consistent excess of the A blood group in classes I and II and a general deficit of A in classes III–V. The numbers of group O individuals are the converse of this and, on the whole, the pooled B+AB group follows the pattern seen with O group individuals.

Contingency χ^2 tests were carried out on the data for the numbers of the A phenotype in the pooled socio-economic groups I+II and III+IV+V. These tests show that the data for migrants are homogeneous ($\chi^2_1 = 1.23$, $P > 0.1$), those for natives are homogeneous ($\chi^2_1 = 1.29$, $P > 0.1$), while within regions there is, for both regions, highly significant heterogeneity between migrants and natives (south-west England $\chi^2_1 = 46.79$, Yorkshire $\chi^2_1 = 19.81$). The causes of this heterogeneity can be seen by reference to Table 2, which shows the distribution of individuals in socio-economic groups I–V in the four sets of donors. It is clear that a substantially higher proportion of individuals of classes I and II is found among migrants than among natives, an observation in line with much published evidence. Another notable point is that the Yorkshire migrants contain a higher proportion of class III (which includes

Table 1 Percentages of blood donors by ABO group and socio-economic class in migrants and natives in south-west England and Yorkshire with signs of the deviations from expected values (B and AB pooled)

Donor set	Blood group	Socio-economic group					Total no.	%
		I	II	III	IV	V		
Migrant SW England	A	55.68+	48.46+	39.74-	45.55+	40.85-	1,135	44.21
	O	34.05-	38.87-	50.29+	43.43-	47.16+	1,142	44.49
	B	5.95-	10.10	6.50	8.26	7.57	202	7.87
	AB	4.32	2.57+	3.47	2.76+	4.42+	88	3.43
	Total no.	185	584	692	472	634	2,567	100
Native SW England	A	70.69+	49.86+	42.58-	45.77+	40.61-	1,040	44.65
	O	22.42-	41.64-	46.13+	42.10-	48.12+	1,034	44.40
	B	5.17-	6.52	7.85	8.45	7.69	179	7.69
	AB	1.72	1.98	3.44	3.68	3.58	76	3.26
	Total no.	58	353	815	544	559	2,329	100
Migrant Yorkshire	A	53.99+	48.31+	39.92-	37.79-	40.03-	909	42.34
	O	37.42-	40.82-	47.78+	47.83+	45.80+	969	45.13
	B	8.59-	7.97	9.44	10.70	9.97	202	9.41
	AB	0.0	2.90	2.86	3.68	4.20	67	3.12
	Total no.	163	414	699	299	572	2,147	100
Native Yorkshire	A	62.19+	43.64+	43.49+	41.92-	37.83-	1,123	42.41
	O	30.49-	42.16-	45.91+	44.83-	49.92+	1,205	45.51
	B	3.66	11.23	7.73	10.71	9.58	246	9.29
	AB	3.66	2.97	2.87	2.54	2.67	74	2.79
	Total no.	82	472	906	551	637	2,648	100
Four sets pooled	A	57.99+	47.45+	41.61-	43.30-	39.80-	4,207	43.41
	O	33.20-	40.70-	47.37+	44.16-	47.79+	4,350	44.89
	B	6.35-	9.22	7.87	9.43	8.70	829	8.55
	AB	2.46	2.63	3.15	3.11	3.71	305	3.15
Grand total		488	1,823	3,112	1,866	2,402	9,691	100

Table 2 Percentage composition, by socio-economic group, of natives and migrants in the two regions

Region	Migrational status	Socio-economic group					n
		I	II	III	IV	V	
SW England	Migrants	7.2	22.7	27.0	18.4	24.7	2,567
	Natives	2.5	15.2	35.0	23.3	24.0	2,329
Yorkshire	Migrants	7.6	19.3	32.6	13.9	26.6	2,147
	Natives	3.1	17.8	34.2	20.8	24.1	2,648

skilled workers) than do the south-west England migrants. This probably reflects the relative concentrations of engineering industry in the two regions. The difference in the proportions of the different socio-economic groups in migrants and natives coupled with the difference in the proportion of the blood groups in different socio-economic groups gives rise to the observed heterogeneity within regions.

However, the qualitative pattern of deviations from expectation is very similar in the four donor sets, and Table 3 shows that the χ^2 value testing the deviations within each set is highly significant. Figure 1 shows the overall effect in the graphical form of the ratio $A/(O+B+AB)$. The data for the two sexes are given separately and the main effect is seen in both sexes, although there are interesting differences between sexes in the two comparisons between socio-economic classes I and II and between IV and V. (The overall distribution of ABO types independent of socio-economic group does not differ significantly between the sexes in the Yorkshire sets but does differ in the sets from south-west England¹³.)

The agreement between two widely separated regions with marked differences in their economic bases, social history and a variety of other factors in the type of socio-genetic differentiation observed is impressive. The consistency of the main effect of excess A among classes I and II and excess O among classes III-V in the four samples of donors argues strongly that these are real deviations likely to be found in comparable samples from other areas in the United Kingdom. Here we note that the distribution of the Rh⁺ and Rh⁻ phenotypes by socio-economic groups does not differ significantly from random expectations (overall data $\chi^2_4 = 6.37$, $P > 0.2$).

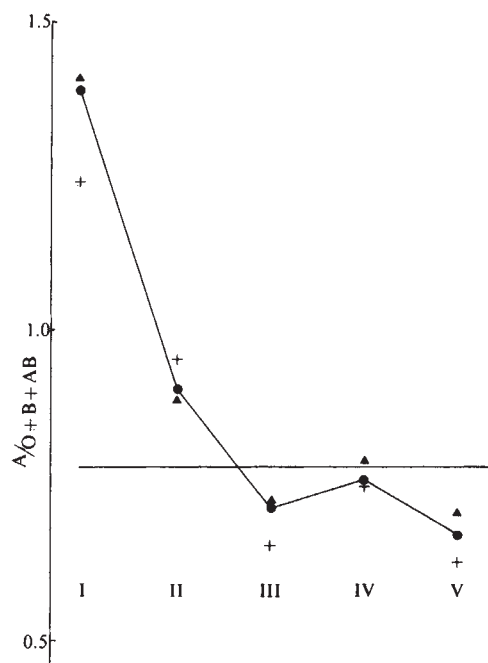


Fig. 1 Ratio of frequency of A to O+B+AB phenotypes in each of socio-economic groups I-V. Data for south-west England and Yorkshire are combined. ●, Sexes combined; ▲, ♂♂; +, ♀♀. The horizontal line gives the ratio expected under random distribution.

The populations under study are composed of blood group donors and it is possible that in some way the procedures attached to collection of blood from donors could produce the observed effect. However, this must be seen only as a formal and probably unlikely cause. As far as biases in collection are known, the most obvious slightly inflates the proportion of the O Rh⁻ type of donors as compared with the general population because of the importance in transfusion of this relatively uncommon type (D. Tills, personal communication). Thus, it is difficult to see how, even if there is over- or under-

Table 3 Summary of contingency χ^2 values and probabilities for distribution of A, O and (B+AB) blood groups in socio-economic groups I-V

Donor set	χ^2	P
Migrants SW England	30.38	<0.001
Natives SW England	27.38	<0.001
Migrants Yorkshire	22.35	<0.005
Natives Yorkshire	24.77	<0.005

representation of particular social classes among blood donors as compared with the general population, the observed effect could arise from blood collection procedures. Mis-classification of blood type, if it does occur, should also not produce a differential effect with respect to socio-economic groupings.

Arguments¹⁻⁴ have been put forward to support the view that the joint effects of partial genetic determination of intelligence, social mobility, assortative mating and the correlation between intelligence and socio-economic class will cause a degree of genetic differentiation in socially stratified human populations.

Gibson *et al.*¹⁰ studied a small population from Otmoor and determined intelligence quotient (IQ; Wechsler) and ABO type. They found differences in IQ in the form $A_2 > O > A_1$. However, the total population size was small (534), it contained related individuals, not all of the seven villages examined showed the differences in IQ and the sample as a whole may not have been representative of the general population. In particular, the natives were greatly outnumbered by non-locally born individuals and it was not possible to determine the effects of migration from other genetically distinct populations.

Dawson¹¹ examined ABO phenotypes in relation to occupational type in 9,867 blood donors in Dublin. No association was detectable but the resolving power of the occupational typing was much less than in the present study so that, for example, "an 'engineer' is placed in the professional class even though he may have been a skilled manual worker" (ref. 11). Mis-classification of this sort would tend to homogenize any initial differences and it is particularly important in Dawson's study as the sample size in class I was appreciably smaller than that in the present study.

A study¹⁴ of two hospital populations in Chile demonstrated a frequency of the A phenotype among patients of a private clinic about 10% higher than among those of a National Health Service hospital. This could suggest a socio-economic differentiation in this population. However, the ethnically highly heterogeneous population of Chile is derived largely from aboriginal Indian people originally monomorphic (or nearly so) for O and from Spanish settlers with a typically European ABO polymorphism. Admixture has been far from random since Europeans entered Chile and private patients have a larger proportion of European ancestry than do National Health Service patients. Thus, any socio-genetic differentiation in this population would inevitably be confounded with the effects of immigration and subsequent nonrandom biological and social mixing.

Many semi-skilled and unskilled individuals have, in the past, migrated from Ireland to the United Kingdom and as Ireland has a high level of the O phenotype, we examined this as a possible cause of the observed distribution. The surnames in samples of individuals from the sets of native donors in both regions were examined in socio-economic groups I, IV and V. Irish names were identified as follows. South-west England: I, 0/40; IV, 2/150; V, 2/130. Yorkshire: I, 2/75; IV, 5/120; V, 0/120. Differential class migration from Ireland cannot therefore be of any significance.

Relatively long-standing differentiation within the population arising from invasions by, or immigration from, genetically well differentiated populations, assortative mating and continuing social stratification as between natives and incomers appears to be an unlikely cause of the distributions we report. The two regions have different invasion and immigration histories and furthermore, social mobility in the United Kingdom has been,

for some generations, a factor which would tend to homogenize initially different gene pools¹⁵.

The only apparently reasonable explanation is that which supposes that the I^A gene when homozygous or when heterozygous with i^o (though not with I^B) confers on those who possess it a probability of ~15% greater than expected of being in social classes I or II. Our data do not permit us to determine the proportions of individuals born into specific classes and those entering later.

This conclusion does not necessarily mean that the ABO genes themselves are responsible for the difference. They may simply be acting as markers for other genes located on the same chromosome pair as the ABO locus, but then only if they are very close and in linkage disequilibrium with it. However, in view of the many associations demonstrated with ABO phenotypes, it does not seem unreasonable to suppose that some of these associations result from pleiotropic effects of the genes and that such effects may include some ultimately expressed as influences on abilities significant in determining the occupational type and hence the socio-economic class of individuals.

Received 25 January; accepted 29 March 1983.

1. Herrnstein, R. J. *I.Q. in the Meritocracy* (Lane, London, 1973).
2. Jensen, A. R. *Educability and Group Differences* (Constable, London, 1973).
3. Thoday, J. M. & Gibson, J. B. *Science* **167**, 990-992 (1970).
4. Halsey, A. H. *Br. J. Sociol.* **9**, 15-28 (1958).
5. Mourant, A. E., Kopeć, A. C. & Domaniewska-Sobczak, K. *The Distribution of the Human Blood Groups and Other Polymorphisms* (Oxford Medical, London, 1976).
6. Kopeć, A. C. *The Distribution of the Blood Groups in the United Kingdom* (Oxford University Press, 1970).
7. Sunderland, E. in *Racial Variation in Man* (ed. Ebling, F. J.) (Institute of Biology, London, 1975).
8. Mourant, A. E., Kopeć, A. C. & Domaniewska-Sobczak, K. *Blood Groups and Diseases* (Oxford Medical, London, 1978).
9. Vogel, F. *Am. J. hum. Genet.* **22**, 464-475 (1970).
10. Gibson, J. B., Harrison, G. A., Clarke, V. A. & Hiorns, R. W. *Nature* **246**, 498-499 (1973).
11. Dawson, G. W. P. *Ann. hum. Genet.* **22**, 315-322 (1958).
12. Registrar-General's Classification of Occupations (HMSO, London, 1966).
13. Fisher, R. A. & Roberts, J. A. F. *Nature* **151**, 640-641 (1943).
14. Valenzuela, C. Y. & Horb, Z. *Social Biol.* **24**, 225-233 (1977).
15. Halsey, A. H. *Change in British Society* (Oxford University Press, 1978).

Benoxaprofen suppression of polychlorinated biphenyl toxicity without alteration of mixed function oxidase function

Arleen B. Rifkind & Haidrun Muschick

Departments of Pharmacology and Medicine, Cornell University Medical College, 1300 York Avenue, New York, New York 10021, USA

Polyhalogenated hydrocarbons are widely distributed environmental pollutants. Several members of the class, including certain polychlorinated biphenyls (PCBs) and polychlorinated dibenzofurans and dibenzodioxins, produce a characteristic toxicity syndrome, manifestations of which are increased mortality, oedema, hyperkeratosis, thymic involution and hepatotoxicity^{1,2}. The toxic hydrocarbons are also inducers of cytochrome P448-mediated mixed function oxidases^{2,3}. The toxicity and induction responses both involve initial binding of the hydrocarbon to the same cytosolic receptor³, but the subsequent events are not understood. It is not known, for example, whether the toxicity and induction are causally related, or whether they are coordinated but independent aspects of a pleiotropic response³. Here we report that benoxaprofen, a nonsteroidal antiinflammatory agent, decreases the toxicity of a PCB isomer, 3,4,3',4' tetrachlorobiphenyl (TCB), in the chick embryo and that it does so without altering the degree of mixed function oxidase induction. The independent alteration of PCB toxicity and induction suggests that the two phenomena are not causally related. The decrease in toxicity by benoxaprofen suggests further that products of arachidonic acid metabolism or other mediators of inflammation may have a causal role in halogenated hydrocarbon toxicity.

Table 1 Effects of 3,4,3',4' tetrachlorobiphenyl and 3,4,3',4' tetrachlorobiphenyl plus benoxaprofen on development of the chick embryo

	Control 6/14	TCB 17/39*	TCB + benoxaprofen 7/41
Intraembryonic death (no. of embryos dead/total treated)			
Embryo weight (g)	14.6 ± 0.4	14.5 ± 0.5	14.6 ± 0.4
Embryo length (mm)	73 ± 0.8	72 ± 0.9	73 ± 0.7
Pericardial oedema (no. with oedema/total live embryos)	1/35	4/22*	1/34
Thymus weight (mg)	115 ± 7	91 ± 6*	99 ± 6
Liver weight (mg)	371 ± 14	434 ± 18*	412 ± 14*
Spleen weight (mg)	9.7 ± 0.4	11.9 ± 0.7*	11.3 ± 0.6*
Bursa weight (mg)	16.2 ± 0.9	14.5 ± 0.9	15.8 ± 1.0
Lung weight (mg)	133 ± 6	126 ± 8	135 ± 6

Chick embryos, White Leghorn strain, were injected at 9 days post-fertilization with 0.02 µl dioxane in two 0.01-µl doses 1 h apart (control); 0.01 µl dioxane followed in 1 h by 60 nmol 3,4,3',4' tetrachlorobiphenyl in 0.01 µl dioxane (TCB); or 0.2 mg benoxaprofen in 0.01 µl dioxane followed in 1 h with 60 nmol TCB in 0.01 µl dioxane (TCB + benoxaprofen). Eggs were injected through a small hole made in the shell using a 1½-inch needle inserted to the hilt, incubated at 37.5 °C and 70% humidity for 9 days and then opened. The number of dead embryos was counted. Live embryos were weighed and measured. Livers, lungs, thymuses, bursas and spleens were removed and weighed. Embryos were inspected for pericardial oedema. Because of the time required for the dissections and measurements a series of six experiments were carried out at successive 2-week intervals using in each experiment six to seven eggs in each treatment group. The data presented were pooled and analysed as one experiment rather than as a series of group means from separate experiments. The results by both types of analyses were essentially the same. The analysis as presented permits the information for each embryo to be weighed equally. For tests of statistical significance, *P* values were determined by Fisher's exact test for incidence of embryonic death and pericardial oedema and by Student's *t*-test for the other data. The number of embryos examined were 41, 39 and 41 for control, TCB and TCB plus benoxaprofen groups, respectively. For indices other than intraembryonic mortality the group sizes reflect the number of live embryos (35, 22 and 36 for each of the three groups, respectively). Means are ± 1 s.e.

* Difference from control values was statistically significant at a *P* value of 0.05 or less.

3,4,3',4' TCB (Ultra Scientific Co., dibenzodioxin and dibenzofuran free), a PCB with demonstrated toxic^{4,5} and P448-inducing⁶ potential, was administered alone or preceded by benoxaprofen (courtesy of Lilly) or indomethacin (Sigma) to chick embryos at 9 days gestational age. The embryos were examined at 18 days gestational age. Table 1 summarizes the findings. The doses of TCB, benoxaprofen and indomethacin were chosen on the basis of exploratory dose-response studies for each agent administered alone. For TCB, mortality was over 70% at 75 nmol TCB per egg, but was not increased by doses of ≤ 20 nmol per egg. A dose of 60 nmol per egg (17.6 µg per egg) was chosen to allow decreases in toxicity to be discerned. At doses of benoxaprofen of ≥ 1 mg per egg, spleen weights were decreased in some embryos. For benoxaprofen, we used a dose of 0.2 mg per egg, well within the range in which no independent effects were observed. Indomethacin was without independent effect at doses up to 3 mg per egg and that dose was used in the experiments reported below.

The experiments summarized in Table 1 show that TCB significantly increased mortality, from 15% in control embryos to 44%. Among the live embryos, embryo length and weight were not affected. The incidence of pericardial oedema was significantly increased. Thus, 18% of embryos had pericardial oedema after TCB treatment, compared with 3% of controls. Liver and spleen weights were significantly increased to 117 and 122% of control weights, respectively, while thymus weights were significantly decreased to 79% of control values.

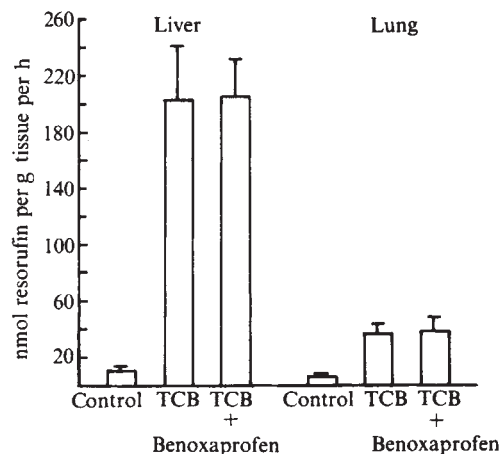


Fig. 1 Effects of 3,4,3',4' tetrachlorobiphenyl and 3,4,3',4' tetrachlorobiphenyl plus benoxaprofen on 7-ethoxyresorufin deethylase activity in chick embryo liver and lung. Livers and lungs from the embryos for which data are shown in Table 1 were assayed individually for 7-ethoxyresorufin deethylase as an index of cytochrome P448-mediated mixed function oxidase activity. Livers and lungs were homogenized in 0.1 M Tris-HCl buffer, pH 8.0 (25% w/v), and then centrifuged at 9,000g for 15 min. The assay procedure was a modification of the procedure of McCormack *et al.*⁸. Incubation mixtures contained NADPH, 1 mM, 7-ethoxyresorufin, 2.0 µM, 2.0 mg wet weight equivalent of 9,000g supernatant for liver and 4 mg for lung, and 0.1 M Tris-HCl, pH 8.0 in a total volume of 0.24 ml. Reaction mixtures were incubated in 6 × 50 mm glass tubes for 10 min at 37 °C in air and reactions were stopped with 0.25 ml cold acetone. The tubes were centrifuged to concentrate the pellet at the bottom. They were placed without decanting in a Hitachi MFP 3 spectrophotofluorimeter to measure fluorescence. Excitation and emission wavelengths were 558 and 590 nm, respectively. A quinine sulphate standard previously standardized against resorufin was used to measure the amount of product formed. Reaction time and tissue concentrations were within the ranges in which product formation was linear for both liver and lung. Benoxaprofen alone did not significantly affect 7-ethoxyresorufin deethylase activity. Means are ± 1 s.e., *n* = 35, 22 and 36, for control, TCB and TCB and benoxaprofen groups, respectively. The values in TCB- and benoxaprofen-treated groups differed significantly from the controls (*P* < 0.001), but did not differ significantly from each other.

Bursa weights were 90% of control values after TCB treatment, but the decrease was not statistically significant. Lung weights were not significantly altered.

As some of the manifestations of TCB toxicity such as oedema and immune system changes resemble symptoms that can accompany inflammatory reactions, it seemed of interest to examine whether an antiinflammatory agent might modify TCB toxicity. Benoxaprofen and indomethacin were used as antiinflammatory agents.

The experiments summarized in Table 1 show that benoxaprofen pretreatment reduced mortality and the incidence of pericardial oedema approximately to control levels. Thymus weights in embryos treated with benoxaprofen and TCB no longer differed significantly from thymus weights in controls. Although liver and spleen weights were closer to normal after benoxaprofen pretreatment the changes were not statistically significant.

7-Ethoxyresorufin deethylase activity was used as an index of cytochrome P448-mediated mixed function oxidase activity^{7,8}. Figure 1 shows the effects of TCB and TCB plus benoxaprofen on 7-ethoxyresorufin deethylase activity in the livers and lungs of embryos which were alive at 19 days gestation and for which other data are shown in Table 1. 7-Ethoxyresorufin deethylase activity in control embryos was slightly higher in liver than in lung. Nine days after a single dose of 60 nmol TCB per egg the deethylase activity was increased 20-fold in the livers and more than 5-fold in the

lungs. 7-Ethoxyresorufin deethylase activities were essentially the same after benoxaprofen plus TCB treatment as after TCB treatment alone.

Indomethacin (protocol as described in Table 1 legend except that indomethacin was used rather than benoxaprofen) did not alter either TCB toxicity or TCB induction of 7-ethoxyresorufin deethylase activity. In embryos treated with TCB (60 nmol per egg) or TCB plus indomethacin (3 mg per egg), mortality was 49% (of 45 embryos) and 51% (of 45 embryos), respectively. Among live embryos, thymus weights were 80 and 73% of control levels respectively, while pericardial oedema was present in 17 and 19% of the embryos, respectively.

The results show that 3,4,3',4' TCB, a PCB that is toxic in adult mammals^{4,5}, is also toxic in chick embryos. Thus, embryos injected with a single low dose of TCB exhibited an increase in mortality and in the incidence of pericardial oedema and a decrease in thymus weight, changes typical of the polyhalogenated hydrocarbon toxicity complex. A decrease in body or bursa weight, as sometimes reported in halogenated hydrocarbon toxicity^{1-3,9}, was not observed in our embryos. Thus, the decrease in thymus weight could not be attributed to an overall growth impairment. Spleen weights were increased in TCB-treated embryos. Decreased spleen weights have sometimes been reported in animals who died or exhibited decreased body weights after halogenated hydrocarbon exposure, but there have also been reports of splenomegaly after sublethal exposures¹⁰.

The degree of toxicity was decreased by benoxaprofen treatment. Some toxic manifestations were improved more than others: mortality principally, followed by pericardial oedema and then thymic involution. It was less clear that benoxaprofen affected the other toxic manifestations. The different degrees of response for the various toxic manifestations could reflect differences in their aetiologies or differences in the degree of access of benoxaprofen to the affected sites.

This is the first example of which we are aware, of protection against halogenated hydrocarbon toxicity by a pharmacological agent. Benoxaprofen has multiple actions, which include inhibition of cyclooxygenase¹¹, inhibition of lipoxygenase¹² and inhibition of chemotaxis¹³. Any of those actions could have contributed to the modulation of TCB toxicity. The results suggest that products of arachidonic acid metabolism or other mediators of inflammation may lead to some of the manifestations of polyhalogenated hydrocarbon toxicity.

The failure of indomethacin but effectiveness of benoxaprofen in decreasing TCB toxicity may signify that lipoxygenase products of arachidonic acid rather than the cyclooxygenase products, the prostaglandins, participate in halogenated hydrocarbon toxicity. That possibility seems to be consistent with the report¹⁴ that tetrachlorodibenzodioxin did not increase prostaglandin synthetase activity in rabbit liver and kidney. However, as the authors noted¹⁴, their findings do not exclude a role for prostaglandins in halogenated hydrocarbon toxicity; neither do our findings with indomethacin. Thus, pharmacokinetic rather than pharmacodynamic factors could account for the differences in the effects of benoxaprofen and indomethacin on TCB toxicity. Indomethacin has a brief plasma half life (20 min to 4 h in various species)¹⁵ and is extensively metabolized by pathways functional in the chick embryo¹⁵. Benoxaprofen, on the other hand, has a long plasma half life (≥ 24 h in rat, mouse and man¹⁶), and virtually all of its metabolism is by glucuronidation, a process that is inactive in the chick embryo before hatching¹⁷. Benoxaprofen is likely to remain largely unmetabolized in the embryo, thereby offering a pharmacokinetic advantage. We may have been able to observe its ameliorative effect because of its sustained actions in our system. We were unable, however, readily to assess the effect of repeated doses of indomethacin on toxicity because the repeated injections required are themselves injurious.

The amelioration of toxicity by benoxaprofen without a coordinate decrease in mixed function oxidase induction indicates that the two phenomena, toxicity and induction, can be altered

independently of each other. Notwithstanding the evidence that the same cytosolic receptor mediates both halogenated hydrocarbon toxicity and P448-type mixed function oxidase induction, there has not yet been any experimental evidence that the mixed function oxidases participate directly in the toxicity³.

The data presented here are compatible with the possibilities either that the induction of mixed function oxidases initiates, through an activated PCB, a secondary series of events leading to toxicity, some of which are blocked by benoxaprofen, or that the induction and toxicity are independent aspects of a coordinated pleiotropic response as suggested by Poland³. The evidence that 3-methylcholanthrene and β -naphthoflavone, cytochrome P448 inducers that differ structurally from TCB, also decrease thymic weights^{18,19} makes the first possibility unlikely. Thus, the data presented here support the hypothesis that the induction and toxicity are coordinated but not causally related events.

This research was supported by NIH grant ES-02165.

Received 6 December 1982; accepted 8 April 1983.

1. Kimbrough, R. *CRC Crit. Rev. Tox.* **2**, 445-498 (1974).
2. Goldstein, J. A. in *Halogenated Biphenyls, Terphenyls, Naphthalenes, Dibenzodioxins and Related Products* (ed. Kimbrough, R.) 151-190 (Elsevier, Amsterdam, 1980).
3. Poland, A. & Knutson, J. C. *Rev. Pharmac. Tox.* **22**, 517-554 (1982).
4. Yoshimura, H., Yoshihara, S., Ozawa, N. & Miki, M. *Ann. N.Y. Acad. Sci.* **320**, 179-192 (1979).
5. Silkworth, J. B. & Grabstein, E. M. *Tox. appl. Pharmac.* **65**, 109-115 (1982).
6. Poland, A. & Glover, E. *Molec. Pharmac.* **9**, 736-747 (1973).
7. Burke, M. D. & Mayer, R. T. *Drug Metab. Disposition* **2**, 583-588 (1974).
8. McCormack, K. M., Cagen, S. Z., Rickert, D. C., Gibson, J. E. & Dent, J. G. *Drug Metab. Disposition* **7**, 252-259 (1979).
9. Shoya, S., Horiuchi, T. & Kohanawa, M. *Natn. Inst. Anim. Hlth Q., Tokyo* **19**, 53-64 (1979).
10. McConnell, E. E. in *Halogenated Biphenyls, Terphenyls, Naphthalenes, Dibenzodioxins and Related Products* (ed. Kimbrough, R.) 109-150 (Elsevier, Amsterdam, 1980).
11. Cashin, C. H., Dawson, W. & Kitchen, E. A. *J. Pharm. Pharmac.* **29**, 330-336 (1977).
12. Walker, J. R. & Dawson, W. *J. Pharm. Pharmac.* **31**, 778-780 (1979).
13. Ford-Hutchinson, A. W., Walker, J. R., Connor, N. S., Oliver, A. M. & Smith, M. J. H. *J. Pharm. Pharmac.* **29**, 372-373 (1977).
14. Kohli, K. K. & Goldstein, J. A. *Life Sci.* **29**, 299-305 (1981).
15. Shen, T.-Y. & Winter, C. A. *Adv. Drug Res.* **12**, 89-245 (1977).
16. Chatfield, D. H. & Green, J. N. *Xenobiotica* **8**, 133-144 (1978).
17. Dutton, G. J. & Ko, V. *Biochem. J.* **99**, 550-556 (1966).
18. Poland, A. & Glover, E. *Molec. Pharmac.* **17**, 86-94 (1980).
19. Blumer, J. L. & Frank, D. M. *Fedn Proc.* **40**, 745 (1981).

Sensitization of Gram-negative bacteria to antibiotics and complement by a nontoxic oligopeptide

Martti Vaara

National Public Health Institute, SF-00280 Helsinki 28, Finland

Timo Vaara

Department of Microbiology, University of Helsinki, SF-00710 Helsinki 71, Finland

A major virulence factor of bacteria that cause generalized infections is their resistance to the lytic action of the complement cascade¹, an important defence mechanism of the host. Invasive Gram-negative enteric bacteria, which cause about one-third of all bacteraemic infections^{2,3}, are completely resistant to lysis by complement, even in the presence of hyperimmune serum⁴⁻⁶. The same bacteria are also resistant to many antibiotics that are effective therapeutic agents against other bacteria, as the outermost surface layer (the outer membrane) of the bacteria functions as a permeability barrier⁷. Here we show that it is possible to sensitize such bacteria to both complement and antibiotics by using an agent that binds to the outer membrane. This agent is a nontoxic derivative of polymyxin which by itself has no bactericidal action^{8,9}.

Polymyxin, a polycationic amphipathic peptide antibiotic, is bactericidal to Gram-negative bacteria and has a dual mechanism of action. First, it binds to the outer membrane¹⁰, leading to structural disorganization of the latter that allows the entry of polymyxin itself and other hydrophobic or amphipathic agents¹¹⁻¹⁴. It then binds to the cytoplasmic membrane, leading

Table 1 Minimum inhibitory concentrations ($\mu\text{g ml}^{-1}$) of hydrophobic antibiotics and penicillin against *E. coli* IH3080 (O18:K1) in the presence of PMBN

	[PMBN] ($\mu\text{g ml}^{-1}$)				
	0	0.3	1	3	10
Erythromycin	30	1	1	1	1
Clindamycin	30	3	3	3	3
Rifampicin	3	3	0.1	0.03	0.03
Fusidic acid	100	3	3	3	1
Novobiocin	30	10	10	3	3
Cloxacillin	>300	100	100	30	30
Penicillin	10	10	10	3	3

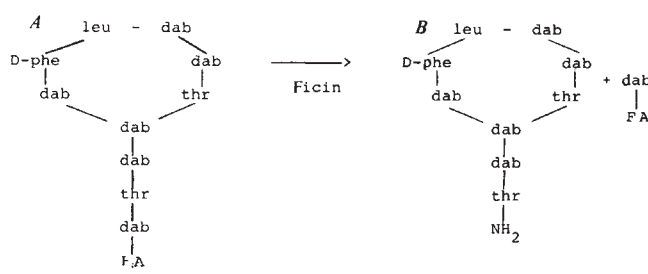
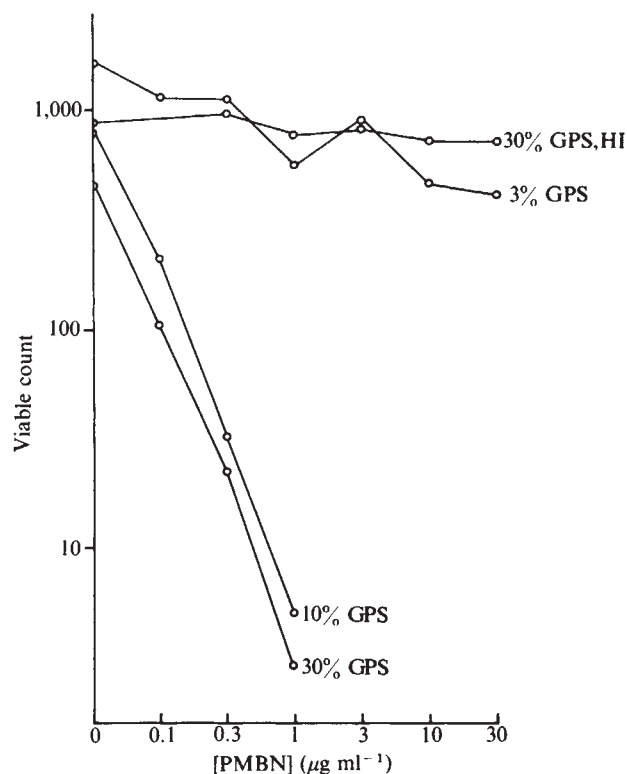
Bacteria (10^4 cells ml^{-1}) were grown in Davis minimal medium²⁴ containing (ml^{-1}): 1 mg glucose, 1 mg 'casaminoacids' (Difco), 20 μg L-tryptophan and increasing amounts of the antibiotic to be tested, as well as of PMBN. Visual growth was recorded after overnight incubation at 37 °C.

to leakage of cytoplasmic components and cell death^{11,15}. Unfortunately, polymyxin also attacks the cytoplasmic membrane of eukaryotic cells¹⁶ and is therefore too toxic to be clinically useful¹⁷. Removal of the fatty acyl group from polymyxin is known to greatly reduce its toxicity; however, the bactericidal activity is lost at the same time^{8,9}. We thought it possible that this modified polymyxin might retain the first action of the parent compound, that is, disorganization of the outer membrane. This would make the membrane permeable to antibiotics and perhaps also vulnerable to attack by the complement cascade.

We removed the fatty acyl group and the terminal diaminobutyric acid residue from polymyxin B by using the proteolytic enzyme ficin and the method of Chihara *et al.*⁸ to obtain polymyxin B nonapeptide (PMBN; Fig. 1). As a test organism we used a smooth encapsulated (O18:K1) *Escherichia coli* strain IH3080 (original strain code EM40; ref. 18) isolated from the cerebrospinal fluid of a human neonate with meningitis. When tested in a growth inhibition test, strain IH3080 was, as expected for an *E. coli* strain, highly resistant to hydrophobic antibiotics (erythromycin, clindamycin, rifampicin, fusidic acid, novobiocin and cloxacillin) which are known to be excluded by the intact outer membrane^{7,19}. Addition of PMBN to the test system sensitized the bacteria to all these agents (Table 1). Concentrations of PMBN as low as 0.3 $\mu\text{g ml}^{-1}$ were sufficient to cause a measurable effect, and the maximum sensitization obtained was 100-fold. Only a slight sensitivity increase was observed to penicillin, a hydrophilic antibiotic which is believed to diffuse across the outer membrane through hydrophilic pores⁷.

We then tested other Gram-negative bacteria which cause severe or bacteraemic infections. Clinical isolates of *E. coli* (serotypes O2:K1, O4:K12, O6:K2, O18:K5 and O75:K5), *Klebsiella pneumoniae*, *Kl. oxytoca*, *Enterobacter agglomerans* and *Salmonella typhimurium* (one strain of each tested) as well as *Pseudomonas aeruginosa* (two strains tested) were as sensitive as IH3080 to the action of PMBN, showing increased sensitivity to hydrophobic antibiotics (erythromycin, fusidic acid and novobiocin). By contrast, inherently polymyxin-resistant strains (*Proteus mirabilis*, two strains) were resistant. Therefore, it seems possible that PMBN could be used clinically in the treatment of Gram-negative infections, to extend the narrow antibacterial spectrum of many hydrophobic antibiotics which are otherwise useful.

Because PMBN increases the permeability of the outer membrane without affecting the viability of the bacteria, it might also be a valuable tool in bacterial research, allowing the entry of hydrophobic mutagens, antimetabolites, inhibitors and other probes into the cells. So far, the only comparable method for this purpose has been treatment with Tris-EDTA²⁰, which, however, has many problems: prolonged treatment (over 5 min) with EDTA is toxic to the bacteria²¹ and after removal of

**Fig. 1** Structural formula of polymyxin B (A) and its derivative PMBN (B), prepared by an enzymatic (ficin; EC 3.4.22.3) treatment as described elsewhere⁸. dab, diaminobutyric acid; phe, phenylalanine; leu, leucine; thr, threonine; FA, fatty acyl group. Polymyxin and ficin, P4125, were from Sigma.**Fig. 2** Sensitivity of *E. coli* IH3080 (O18:K1) to fresh guinea pig serum (GPS) in the presence of PMBN. Bacteria (300 cells ml^{-1} , grown to early logarithmic phase in L-broth²⁶) were incubated in whole serum or the indicated dilutions of serum (in phosphate-buffered saline) in the presence of PMBN for 2 h at 37 °C, then the number of viable bacteria for each dilution was determined by plating onto L-plates (values are related to the number 100 at the beginning of the experiment). In the HI serum, the complement was inactivated by heating at 56 °C for 30 min.

EDTA, the cells rapidly restore their outer membrane permeability barrier²⁰.

Next we tested the possible effect of PMBN on the interaction of the bacteria with complement. In the absence of PMBN, IH3080 was, as expected, completely resistant to normal serum complement, surviving and even multiplying in 30% fresh guinea pig serum. However, concentrations of PMBN as low as 0.3 $\mu\text{g ml}^{-1}$ sensitized the bacteria to the bactericidal action of both 30% and 10% guinea pig serum (Fig. 2). The bactericidal action of serum plus PMBN was abolished if the serum was heated to inactivate complement. Thus, PMBN alone was not bactericidal but sensitized the bacteria to the action of complement in the serum. In comparable experiments, four other smooth encapsulated serotypes of *E. coli* (O2:K1, O4:K12, O18:K5 and O75:K5) and a smooth *S. typhimurium* were tested; PMBN rendered all of them sensitive to 30% guinea pig serum. When guinea pig serum was absorbed in the

Table 2 Both antibodies and PMBN are required for the bactericidal action of guinea pig serum

Additive	30% Absorbed fresh guinea pig serum		30% Absorbed heat-inactivated guinea pig serum
	-PMBN	+PMBN (10 µg ml ⁻¹)	
None	1,500	360	1,500
Normal rabbit serum			
10 ⁻¹	730	<3	1,500
10 ⁻²	720	32	1,500
10 ⁻³	1,500	160	1,500
Rabbit anti-018 hyperimmune serum			
10 ⁻¹	120	<3	1,500
10 ⁻²	200	<3	1,500
10 ⁻³	700	<3	1,500
10 ⁻⁴	1,500	40	1,500
10 ⁻⁵	1,500	220	1,500

Guinea pig serum was absorbed with a mixture (ml⁻¹) of 10¹⁰ cells of glutaraldehyde-fixed IH3080, its 018:K1 derivative, EH817, and its 018:K1 derivatives EH778 and EH786, for 90 min at 4°C to remove anti-IH3080 antibodies. Absorption did not reduce the complement level (haemolytic complement, titrated in a sensitized sheep red blood cell assay²⁵, was unchanged). The indicated dilutions of rabbit sera were then added to the absorbed guinea pig serum and the bactericidal assay using IH3080 was performed as described in Fig. 2 legend. The values indicate the numbers of the bacteria (relative to the number 100 at the beginning of the experiment) that were viable after a 2-h treatment. The complement in the rabbit sera used was destroyed by heating at 56°C for 30 min.

cold to remove possible natural antibodies directed against the surface structures of IH3080, it lost most of its PMBN-dependent bactericidal activity (Table 2). The bactericidal system could be reconstituted by adding to the absorbed serum heated normal rabbit serum (up to a dilution of 10⁻²) or rabbit anti-018 hyperimmune serum (up to a dilution of 10⁻⁴). Table 2 again shows that PMBN was necessary for the bactericidal action: even a high titre of anti-018 hyperimmune serum plus complement was not bactericidal in the absence of PMBN. Thus, complement, PMBN and antibodies are all essential components in the bactericidal system but normal serum is sufficient as the antibody source.

To account for the complement-sensitizing action of PMBN, we propose that its interaction with the outer membrane of the bacteria exposes normally hidden hydrophobic aspects of the membrane, thereby facilitating the insertion of the hydrophobic membrane attack complex²² of complement. Recent evidence indicates that the resistance of smooth *Salmonella* to complement is due to the failure of this complex to insert into the outer membrane while the more hydrophobic surface of rough mutants allows both insertion and killing^{6,23}. Thus, the same mechanism, disorganization of the outer membrane resulting in exposure of a hydrophobic milieu, could at the same time explain the increased permeability of the outer membrane to the hydrophobic antibiotics and the sensitization to killing by complement.

PMBN might offer a novel means of inhibiting enteric bacteria. PMBN by itself is not bactericidal but renders the bacteria vulnerable to various environmental noxa by destroying the main defence mechanisms of the outer membrane. We are now attempting to determine to what extent this principle could be applied to the treatment of bacterial infections.

We thank P. H. Mäkelä for discussions and for critically reading the manuscript, and P. Viljanen for technical assistance. This study was supported by a research grant from the Emil Aaltonen Foundation (to M.V.).

Received 13 January; accepted 8 April 1983.

1. Fine, D. P. *Complement and Infectious Diseases*, 157 (CRC, Boca Raton, Florida, 1981).
2. McCowan, J. E. Jr, Barnes, M. W. & Finland, M. *J. infect. Dis.* **132**, 316-334 (1975).
3. Williams, G. T., Houang, E. T., Shaw, E. J. & Tabagchali, S. *Lancet* **ii**, 1291-1293 (1976).
4. Roantree, R. J. & Ranz, L. A. *J. clin. Invest.* **39**, 72-81 (1960).

5. Roantree, R. J. & Poppas, N. C. *J. clin. Invest.* **39**, 82-91 (1960).
6. Joiner, K. A., Hammer, C. H., Brown, E. J., Cole, R. J. & Frank, M. M. *J. exp. Med.* **155**, 797-808 (1982).
7. Nikaido, H. & Nakae, T. *Adv. microb. Physiol.* **20**, 163-250 (1979).
8. Chihara, S., Tobita, T., Yahata, M., Ito, A. & Koyama, Y. *Agric. Biol. Chem.* **37**, 2455-2463 (1973).
9. Chihara, S., Ito, A., Yahata, M., Tobita, T. & Koyama, Y. *Agric. Biol. Chem.* **38**, 521-529 (1974).
10. Teuber, M. & Bader, J. *Archs Microbiol.* **109**, 51-58 (1976).
11. Teuber, M. *Archs Microbiol.* **100**, 131-144 (1974).
12. Rosenthal, K. & Storm, D. *J. Antibiot.* **30**, 1087-1092 (1977).
13. Vaara, M. & Vaara, T. *Antimicrob. Ag. Chemother.* **19**, 578-583 (1981).
14. Vaara, M. *J. Bact.* **148**, 426-434 (1981).
15. Storm, D. E., Rosenthal, K. S. & Swanson, P. E. A. *Rev. Biochem.* **46**, 723-763 (1977).
16. Kunin, C. M. *J. infect. Dis.* **121**, 55-65 (1960).
17. Garrod, L. P., Lambert, H. P. & O'Grady, F. *Antibiotic and Chemotherapy* (Churchill Livingstone, London, 1981).
18. Paakkanen, J., Gotschlich, E. C. & Mäkelä, P. H. *J. Bact.* **139**, 835-841 (1979).
19. Vaara, M. thesis, Univ. Helsinki, (1981).
20. Leive, L. *Ann. N.Y. Acad. Sci.* **235**, 109-127 (1974).
21. Leive, L. & Kollin, V. *Biochem. biophys. Res. Commun.* **28**, 229-236 (1967).
22. Podack, E. R. & Müller-Eberhard, H. J. *J. Immun.* **121**, 1025-1030 (1978).
23. Joiner, K. A., Hammer, C. H., Brown, E. J. & Frank, M. M. *J. exp. Med.* **155**, 809-819 (1982).
24. Sanderson, K. E., Ross, H., Ziegler, L. & Mäkelä, P. H. *Bact. Rev.* **36**, 608-637 (1972).
25. Keleti, G. & Lederer, W. H. *Handbook of Micromethods for the Biological Sciences* (Van Nostrand, New York, 1974).
26. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).

Basophil variants with impaired cromoglycate binding do not respond to an immunological degranulation stimulus

Nachman Mazurek, Pnina Bashkin, Avigdor Petrank & Israel Pecht

Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot, Israel

Cromoglycate, which inhibits IgE-mediated degranulation of mast cells^{1,2} and a basophilic rat tumour cell line (RBL-2H3) (ref. 3), is a drug widely used in the treatment of allergic asthma. We have demonstrated the presence of specific binding sites for that drug on the membranes of basophils and mast cells⁴ and more recently, we have succeeded in isolating the cromoglycate-binding protein from the membranes of RBL-2H3 cells⁵. These findings together with the chelation by cromoglycate of alkaline-earth ions in low polarity medium^{6,7}, suggest that the binding site of the drug may either be part, or closely related to the calcium gate. To further investigate this, we have selected variants of the RBL-2H3 cell line that are defective in cromoglycate binding but bind IgE normally. In these variants, which have similar histamine content to the parental cells, IgE-mediated challenge did not lead to Ca²⁺ influx, degranulation and histamine release. In contrast, these cells were able to release histamine on exposure to the Ca²⁺ ionophore A23187, indicating that the degranulation mechanism distal to the Ca²⁺ gating step is unaffected. Taken together, these findings suggest that cromoglycate-binding protein has a role in the transmembrane calcium influx which induces degranulation.

Selections for RBL-2H3 cell line variants with normal IgE binding but deficient in binding of cromoglycate (the disodium salt of 1,3 bis-(2-carboxychromon-5-yloxy)-2-hydroxypropane were performed using the fluorescence activated cell sorter (FACS II, Becton-Dickinson). To detect such variants in a dividing basophil cell population, we have double labelled the cells with bovine serum albumin (BSA) covalently conjugated with both cromoglycate and fluorescein (Fl-DSCG-BSA), and with rhodaminated IgE (Rh-IgE). By analysing the degree of staining of a normal suspension of the RBL-2H3 cells, using the cell sorter, we have found that all of the RBL-2H3 cells could bind Rh-IgE, whereas 2-3% of these cells did not bind Fl-DSCG-BSA as can be seen in Fig. 1. To exclude nonspecific absorption of BSA, a parental population of RBL-2H3 cells as well as the cloned variants were routinely reacted with

Fig. 1 Fluorescence histograms representing RBL-2H3 cells double labelled with fluoresceinated DSCG-BSA and rhodaminated monoclonal IgE. (The DSCG-BSA was prepared as described elsewhere⁴ and carried an average of four cromoglycate groups per BSA molecule. The DSCG-BSA conjugate was reacted with fluorescein isothiocyanate in 0.05 M borate buffer pH 9.3. Rh-IgE was prepared by reacting monoclonal IgE specific for DNP hapten with rhodamine isothiocyanate by the same procedure.) 2×10^6 cells were double-labelled by simultaneous incubation in 0.5 ml Tyrode buffer, containing 1% BSA with 40 μ g FI-DSCG-BSA and 40 μ g Rh-IgE for 1 h at 4°C. Cells were then washed twice in Tyrode buffer containing 1% BSA and resuspended in 1 ml Tyrode buffer. The double-labelling fluorescence histograms were obtained on the FACS. The 514.5 nm line from a 400 mW argon ion laser (Model 164, Spectra Physics) was used for excitation. Two photomultipliers were used, PMT 1 for the fluorescein emission with the 540 nm long-pass filter and 560 nm short-pass (Ditric Optics) and PMT 2 for the rhodamine emission with 580 nm long-pass filter (No. 2-73 Corning Glass). Channels 1-506 were collecting signals from PMT 1 and 507-1023 signals from PMT 2. The spillover of either rhodamine signals into fluorescein, or vice versa was minimized using the compensation unit. Peak I, cells reacted with FI-BSA. Peak II, double staining of the parental RBL-2H3 cells. A, B, the windows chosen for sorting out the subpopulation not stained with fluorescein and positive for rhodamine staining (as shown in insert). Peak III subpopulation collected after sorting (negative for FI-DSCG-BSA, but positive for Rh-IgE).

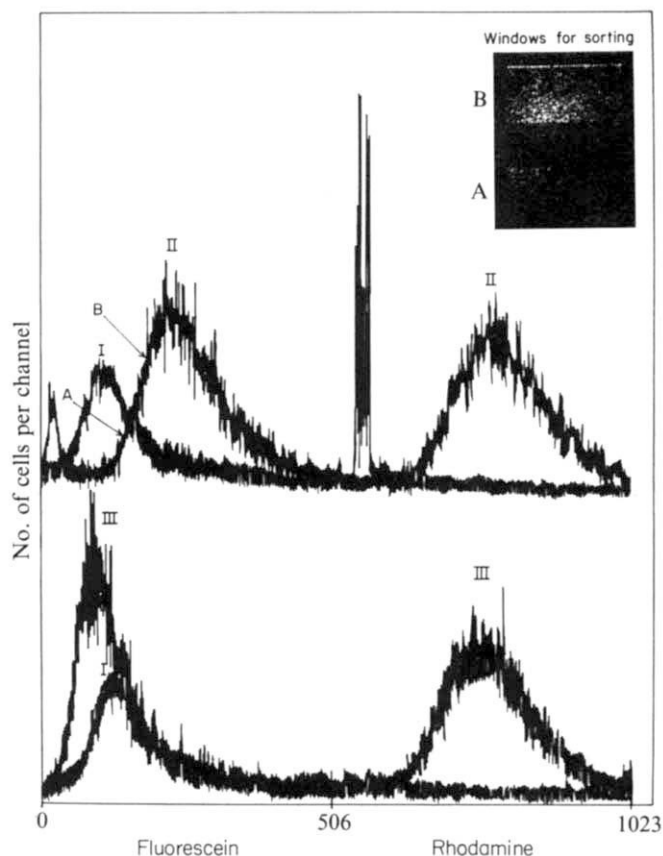
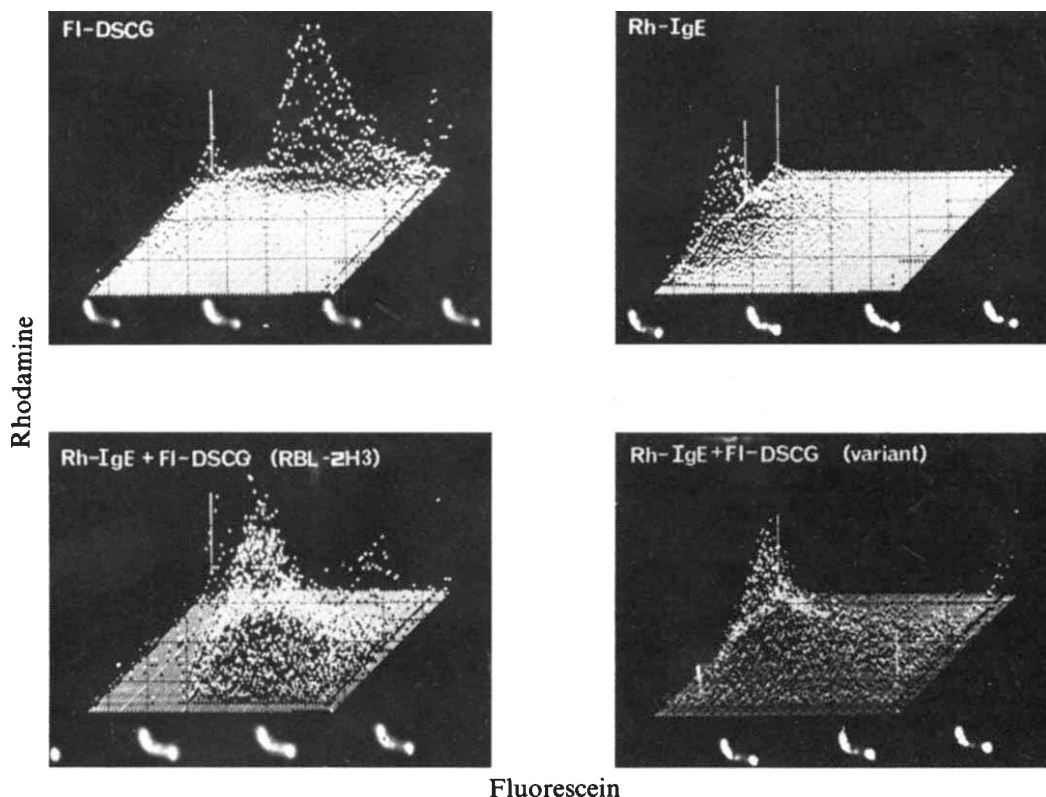


Fig. 2 Three-dimensional presentation of the histograms for the double-labelled parental RBL-2H3 cells and a variant clone RBL-2H3-E7-14. Labelling was performed as for Fig. 1. RBL-2H3 cells could be labelled with fluorescein-DSCG-BSA (top left) or alternatively with rhodaminated IgE (top right). Double labelling of the parental RBL-2H3 cells with both FI-DSCG-BSA and Rh-IgE is shown at bottom left. Similar labelling of the variant cells (bottom right) shows clearly the diminished FI-DSCG binding capacity, whereas the staining with Rh-IgE is unimpaired.



fluoresceinated BSA (FI-BSA). The level of staining with FI-BSA was markedly lower than that observed with specific labelling with FI-BSA-DSCG, as was observed in FACS analysis and remained within the level of the autofluorescence of the cells (Fig. 1, peak I). The resultant subpopulation of IgE receptor positive and DSCG receptor negative cells, was separated by the cell sorter in sterile conditions, cloned by limiting dilu-

tions and recloned on soft agar. Five different clones defective in cromoglycate binding but fully competent of binding IgE were selected (Fig. 2). The number of IgE receptors on these variants is the same as that of the parental cells (determined by radioimmunoassay using ¹²⁵I-IgE). The IgE receptors could be clustered in a manner similar to that observed with the RBL-2H3 cells, using the rhodaminated monoclonal IgE

Table 1 Histamine release and calcium influx in parental and cromoglycate binding defective variant cell lines

Cell line	Net histamine release (%)		Ca ²⁺ influx, pmol per 30 min per 10 ⁶ cells		
	IgE-mediated	Ionophore-induced	Unstimulated	IgE-stimulated	Ionophore-induced
Parental RBL-2H3	21	43	20.5	89.5	93.2
Variants RBL-2H3					
E7-10	0	54	21.5	21.7	96.8
B-5	0	46	23.0	21.4	89.4
E7-14	2	61	22.4	22.0	97.2

The RBL-2H3 cells were grown as a monolayer in 35 mm-diameter plates by plating 10⁶ cells in 3.0 ml of medium. After 48 h, the cells were reacted with 20 µg of 2,4-dinitrophenyl-specific monoclonal IgE for 1 h at 4 °C. Unbound IgE was removed by several washings with 10 mM HEPES-buffered saline. The cells were then incubated for 10 min in 1.0 ml of HEPES-buffered saline (pH 8.0) containing ⁴⁵CaCl₂ at 10 µCi/ml⁻¹ in the presence of 1 mM CaCl₂ (final specific activity 10 Ci mol⁻¹), and then stimulated for histamine release with either 10 µg ml⁻¹ (DNP)₈-BSA or 0.5 µM of the Ca²⁺-ionophore A23187 for 30 min at 37 °C^{10,12}. The medium was removed and assayed for histamine. ⁴⁵Ca²⁺ influx was stopped by washing the attached cells three times with ice-cold saline containing 10 mM HEPES (pH 7.4) and 10 mM CaCl₂ to remove extracellular ⁴⁵Ca²⁺. One ml of a hypotonic 5 mM CaCl₂ solution was then added and the cells were lysed by freezing and thawing. The thawed lysate was collected and centrifuged to remove cellular debris. The radioactivity of each aliquot was measured by liquid scintillation. The data presented are the average of three independent experiments done in duplicates. Cromoglycate at 100 µM concentration was found to inhibit at least 60% of the IgE-mediated degranulation of the RBL-2H3 cells³. To exclude lytic leakage of histamine from the basophils, possibly caused by A23187, we have measured the lactate dehydrogenase (LDH) activity as a monitor for cytoplasmic leakage¹³. Over an ionophore concentration range of 0.25–1 µM, less than 4% increase in LDH activity above basal levels in the medium of both parental and variant cells was observed.

molecules specific to 2,4-dinitrophenyl (DNP) hapten⁸ and either a multivalent DNP-BSA conjugate or anti-IgE antibodies (data not shown). Notably, these variants grew much more slowly than their parental RBL-2H3 cells. Thus their generation time was 40–48 h compared with 18–24 h for the RBL-2H3 cells.

To investigate the possible role of cromoglycate-binding protein in the degranulation process, the ability of the variant cells to allow Ca²⁺ influx and to degranulate was tested. The variant cells were treated with monoclonal IgE specific for the DNP hapten and then cross-linking was induced by treatment with a multivalent DNP-BSA derivative. Alternatively, we exposed the variants to the calcium ionophore, A23187, known to be an effective degranulating agent^{9,10}. We found the total histamine content of the variants to be identical to that of the parental cells (0.2 µg per 10⁶ cells, as measured by the fluorometric assay of Shore *et al.*¹¹). However, no ⁴⁵Ca²⁺ uptake or histamine release by these variants could be induced in the IgE-mediated reaction. Significantly however, these cells underwent degranulation on exposure to A23187, which bypasses the normal initial sequence of membranal events (Table 1). These results indicate that the degranulation mechanism, distal to calcium entry, remained intact in the variants and their inability to degranulate is due to defects in surface events following IgE clustering. Furthermore, these results imply that the cromoglycate-binding protein is involved in the sequence of events triggered by clustering of IgE receptors and might have a key role in transmembrane Ca²⁺ fluxes. But there might be other deficiencies in the variant cells responsible for their inability to undergo IgE-mediated degranulation. For example, RBL cell line-variants deficient in phospholipid methyltransferase I or II were reported also to be impaired in their Ca²⁺ uptake and histamine release capacities¹². Although we did not analyse whether the variants described here are defective in those enzymes, experiments to incorporate the isolated cromoglycate-binding protein into the RBL-2H3-E7-14 variant cells resulted in reconstitution of the Ca²⁺ influx and histamine release³, supporting the notion that the behaviour of the variants is due to the deficiency in the cromoglycate-binding component.

We thank Dr F. Nauwelaers and Mr A. Goldfisher for help with the FACS II, Drs H. Metzger and T. Jovin for the RBL-2H3 cell line, and Z. Eshhar for his helpful discussions.

Received 5 January; accepted 23 March 1983.

- Foreman, J. C. & Garland, L. G. *Br. med. J.* **1**, 820–821 (1976).
- Spataro, A. C. & Bormann, H. B. *Chem. Pharmac.* **25**, 505–510 (1976).
- Mazurek, N., Bashkin, P., Loyter, A. & Pecht, I. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Mazurek, N., Berger, G. & Pecht, I. *Nature* **286**, 722–723 (1980).
- Mazurek, N., Bashkin, P. & Pecht, I. *EMBO J.* **1**, 505–590 (1982).
- Spitzer, J. C. *Int. Archs. Allergy appl. Immun.* **60**, 356–358 (1979).
- Mazurek, N., Geller-Bernstein, C. & Pecht, I. *FEBS Lett.* **111**, 194–196 (1980).
- Eshhar, Z., Ofarim, M. & Waks, T. *J. Immun.* **124**, 69–695 (1980).

- Fewtrell, C. & Metzger, H. *Proc. FASEB Ser.* **14**, 295–314 (1981).
- Fewtrell, C., Lagunoff, D. & Metzger, H. *Biochim. biophys. Acta* **644**, 363–368 (1981).
- Shore, A. P., Buckhalter, A. & Cohen, H. V. *J. Pharmac. exp. Ther.* **127**, 182–186 (1950).
- McGivney, A., Crews, F. T., Hirata, F., Axelrod, J. & Siraganian, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6176–6180 (1981).
- Taurog, J. D., Fewtrell, C. & Becker, E. L. *J. Immun.* **122**, 2150–2153.

Detection of factor VIII/coagulant antigen in human liver tissue

H. V. Stel*, Th. H. van der Kwast† & E. C. I. Veerman*

*Central Laboratory of the Netherlands Red Cross Blood Transfusion Service, Plesmanlaan 125, Amsterdam, The Netherlands
†Department of Pathology, Erasmus University, Rotterdam, The Netherlands

Factor VIII, a high molecular weight glycoprotein complex which has an important role in haemostasis, consists of two immunologically as well as functionally discernible moieties that can be isolated separately. These are factor VIII/von Willebrand factor (FVIII/vWF) which is associated with the factor VIII-related antigen (FVIII_RAg), and the factor VIII/procoagulant activity (FVIII/C) which is associated with the factor VIII/procoagulant antigen (FVIII/C_{Ag}). The FVIII/C activity is decreased or absent in patients with haemophilia A (for a review of the structure and function of the factor VIII complex, see refs 1 and 2). Immunological techniques, combined with cell culture, have demonstrated that FVIII_RAg is present^{3,4} in and synthesized by endothelial cells⁵ and megakaryocytes⁶. However, the organ and/or cell type responsible for the production of FVIII/C has not been established. Indirect evidence derived from organ transplantation in experimental animals suggests that the liver is the most likely organ for FVIII/C production⁷. Here we have used a monoclonal antibody against FVIII/C_{Ag} in combination with a sensitive immunostaining technique to demonstrate the presence of FVIII/C_{Ag} in hepatic sinusoidal endothelial cells.

The production and characterization of monoclonal antibodies against human factor VIII complex will be described elsewhere (H.V.S., in preparation). Briefly, BALB/c mice were immunized with purified factor VIII complex⁸. Lymphocyte hybridization was performed by the method of Galfré *et al.*⁹.

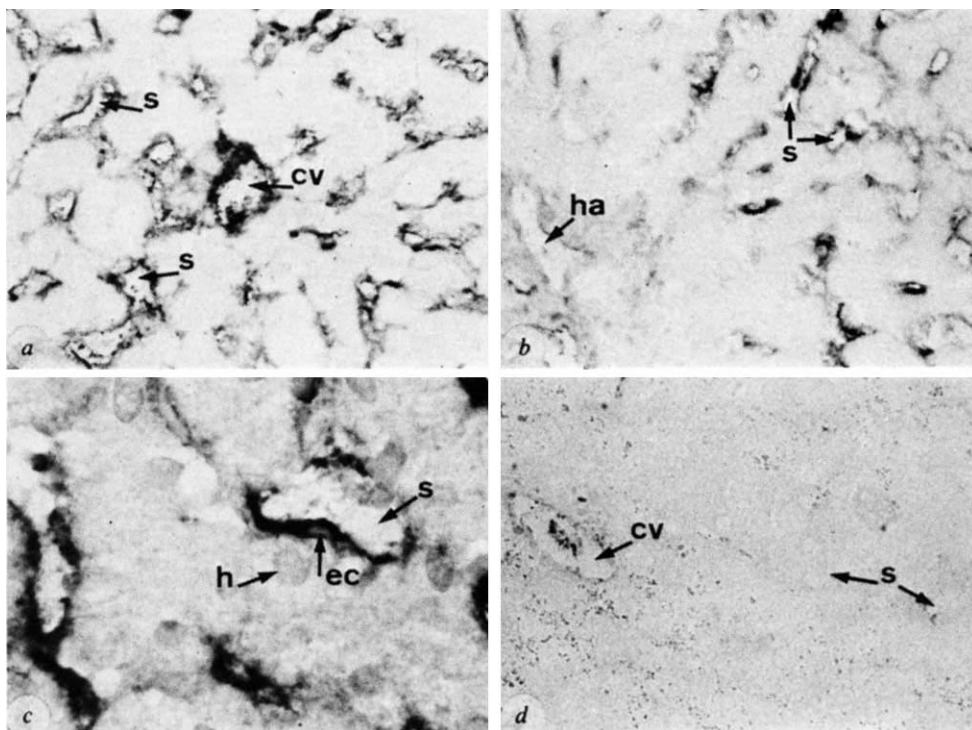
In the study reported here, we used monoclonal antibodies CLB-Rag 1, CLB-Rag 7, CLB-Rag 35 and CLB-Cag A; these are characterized in Table 1. CLB-Rag 1, CLB-Rag 7 and CLB-Rag 35 are directed against FVIII_RAg as judged by the following criteria: (1) binding to purified factor VIII complex, factor VIII present in normal pooled plasma as well as

Table 1 Reactions of the various monoclonal antibodies used

Monoclonal antibody	Binding to:				Inhibition of:	
	FVIII	FVIIIINP	FVIIIHAP	FVIII/CAG	FVIII/C	FVIIIIRCo
CLB-RAg1	+	+	+	-	-	-
CLB-RAg 7	+	+	+	-	-	+
CLB-RAg 35	+	+	+	-	-	+
CLB-CAG A	+	+	-	+	+	-

Binding of monoclonal antibodies to: FVIII, purified factor VIII complex⁸; FVIIIINP, factor VIII present in normal pooled plasma; FVIIIHAP, factor VIII present in haemophilic A plasma (Negative for material cross-reacting with a human antibody against FVIII/CAG as tested in an immunoradiometric assay¹⁴); and FVIII/CAG, factor VIII/CAG¹¹, tested using an enzyme-linked immunosorbent assay (ELISA)¹⁵. For the ELISA, purified factor VIII complex was directly coated onto the wells of the ELISA plates. FVIIIINP and FVIIIHAP were indirectly coated onto the wells of the plates with a rabbit antiserum against factor VIII¹⁴, and FVIII/CAG was indirectly coated onto the wells of the plates with a human antiserum against FVIII/C¹⁶. Inhibition of FVIII/C was tested as described by Muller *et al.*¹⁷. Inhibition of FVIIIIRCo (ristocetin cofactor activity) was tested in a system of ristocetin-induced platelet aggregation¹⁸. + Indicates a positive response in the ELISA (binding three times greater than that of a control mouse monoclonal antibody), inhibition of FVIII/C, or inhibition of FVIIIIRCo, respectively.

Fig. 1 *a*, Cryostat section of liver stained with monoclonal antibody CLB-RAg 1; a diffuse and granular staining is observed in the intima of the central vein as well as in the sinusoidal lining cells. S, sinusoid; CV, central vein. *b*, Cryostat section of the same liver biopsy material as in *a*, but labelled with monoclonal antibody CLB-CAG A (ascitic fluid diluted 1:1,000 in phosphate-buffered saline (PBS)-gelatine, 0.2%). Cells lining the sinusoids are stained. Note that the intima of the hepatic artery (ha) is negative. *c*, Higher magnification of cryostat section of the same liver biopsy material as in *a*, labelled with monoclonal antibody CLB-CAG A and counterstained with Mayer's haematoxylin for 1 min. Endothelial cells lining the sinusoids show strong cytoplasmic staining. Generally, perinuclear staining was seen. Hepatic parenchyma cells showed no deposition of the colour product. h, Nucleus of hepatic parenchyma cells; ec, nucleus of sinusoidal endothelial cell. *d*, Cryostat section of liver biopsy material incubated with control ascitic fluid (diluted 1:1,000 in PBS-gelatine, 0.2%). No immunostaining is observed.



Methods: Cryostat sections (5 μ m thick) of liver were air-dried, fixed for 10 min in acetone and washed in PBS for 10 min. Monoclonal anti-factor VIII undiluted hybridoma supernatant or monoclonal ascitic fluid (diluted 1:1,000 in PBS-gelatine, 0.2%) was overlaid on the tissue and allowed to incubate for 2 h at ambient temperature. Then, immunostaining was performed using a slightly modified immunoperoxidase-staining procedure. The colour reaction was started by adding a PBS solution containing 1 mg ml⁻¹ 3,3'-diaminobenzidine (Sigma) and 0.03% hydrogen peroxide. After incubation for 10 min in the dark at room temperature, the reaction was abrogated by washing in PBS for 30 min.

factor VIII present in haemophilic A plasma, whereas none of these antibodies react with the plasma of a severely affected patient with classical von Willebrand's disease (FVIIIIRAg undetectable); (2) partial (CLB-RAg 1 and CLB-RAg 7) or complete (CLB-RAg 35) inhibition of platelet adherence to human subendothelium tested in a perfusion system, as described by Sakariassen *et al.*¹⁰; and (3) partial (CLB-RAg 7) or complete (CLB-RAg 35) inhibition of ristocetin-induced platelet aggregation.

Evidence that monoclonal antibody CLB-CAG A reacts specifically with FVIII/CAG is provided by the following data (see Table 1): (1) CLB-CAG A binds to purified factor VIII complex as well as to factor VIII present in normal pooled plasma, but not to factor VIII present in three different haemophilic A plasmas (negative for material cross-reacting with a human antibody against FVIII/CAG); (2) FVIII/CAG¹¹ binds to monoclonal antibody CLB-CAG A, but not to CLB-RAg 1, CLB-RAg 7 or CLB-RAg 35; (3) CLB-CAG A inhibits

FVIII/C activity completely; and (4) an immunoadsorbent prepared by covalent linking of monoclonal CLB-CAG A to Sepharose removes FVIII/C and FVIII/CAG, but not FVIIIIRAg from a solution containing the dissociated factor VIII complex. None of these monoclonal antibodies showed binding to human fibrinogen, fibronectin, IgG or albumin.

Immunohistological studies were performed on snap-frozen postmortem material and liver biopsies from 10 patients with mild (chronic persistent hepatitis) or no detectable liver disease. The various monoclonal antibodies were tested on cryostat sections of these tissues with a slightly modified immunoperoxidase-staining procedure¹² (see Fig. 1 legend).

When monoclonal antibody CLB-RAg 1 was tested on liver biopsy material (Fig. 1*a*), the vascular endothelial cells in both large and small hepatic arteries and veins showed immunoperoxidase staining. Immunoreactive material was also noted in the subendothelium around hepatic arteries and veins and in the endothelial cells, and possibly also Kupffer cells, lining

the hepatic sinusoidal spaces. Both diffuse and granular immunoperoxidase deposits were observed in the cytoplasm of the endothelial cells lining arteries, veins and liver sinusoids. Essentially similar results were obtained with CLB-RAg 7 and CLB-RAg 35 (data not shown).

With monoclonal antibody CLB-CAG A, the endothelial cells, and possibly also Kupffer cells, lining the hepatic sinusoidal spaces were positive (Fig. 1b): the cytoplasm of these cells showed diffuse distribution with superimposed coarse granular depositions of the peroxidase reaction product (Fig. 1c). The vascular endothelial cells of the hepatic arteries and veins showed no reaction with CLB-CAG A. When immunostaining was performed with antibody CLB-CAG A which had previously been absorbed with normal pooled plasma, labelling was abolished, whereas absorption of antibody CLB-CAG A with haemophilic A plasma did not affect the results of the immunostaining. In control experiments, performed simultaneously on the same biopsy material, various mouse monoclonal antibodies directed against unrelated antigens yielded negative results (Fig. 1d).

Similar immunoperoxidase-staining experiments were performed using monoclonal antibodies CLB-CAG A and CLB-RAg 1 on liver biopsy material from a patient with severe haemophilia A (FVIII/C < 0.01 U ml⁻¹, FVIII/CAG < 0.01 U ml⁻¹) with chronic persistent hepatitis (this biopsy material was provided by Professor P. M. Mannucci.) No immunostaining was observed with CLB-CAG A. However, when immunostaining was performed using CLB-RAg 1, both diffuse and granular immunoperoxidase deposits were present in the cytoplasm of the endothelial cells lining arteries, veins and liver sinusoids.

The most likely explanation for the observed binding of monoclonal antibody CLB-CAG A is that FVIII/CAG is present in the stained cells, which line the liver sinusoids. After examination of several cryostat sections, it seems to us that the staining pattern is characteristic for sinusoidal endothelial cells. However, it cannot be excluded that Kupffer cells also contain FVIII/CAG. Nonspecific adsorption of the antigen onto the cell surface is not likely, as (1) the immunostaining material is clearly present in the cytoplasm of the cells (Fig. 1c); (2) tissue of other highly vascular organs, such as spleen and kidney, did not show any endothelial immunostaining with monoclonal antibody CLB-CAG A (not shown); and (3) in liver biopsy material of a patient with severe haemophilia A, no FVIII/CAG could be demonstrated, although the FVIII/C plasma level was normalized (1 U ml⁻¹) by administration of factor VIII concentrates before biopsy.

Although immunohistological studies cannot prove directly that FVIII/CAG is synthesized by hepatic sinusoidal endothelial cells, the observation of perinuclear staining with granular depositions of the peroxidase reaction product in the cytoplasm suggests that FVIII/CAG is synthesized in these cells. This is corroborated by Shaw *et al.*¹³, who showed, using a perfusion system, that FVIII/C activity was released from the perfused rat liver. This release was affected by cycloheximide and actinomycin D, which inhibit the *de novo* protein synthesis.

Both FVIII/CAG and FVIIIIRAg are present in the hepatic sinusoidal endothelial cell, which suggests that the whole factor VIII complex is synthesized in this cell. FVIIIIRAg but not FVIII/CAG is synthesized in the vascular endothelial cell⁵. Apparently, two pools of factor VIII are present *in vivo*: one pool consisting of FVIIIIRAg synthesized in vascular endothelial cells and one consisting of the total factor VIII complex synthesized in hepatic sinusoidal endothelial cells.

The results presented here could be important for deducing a means of producing factor VIII by recombinant DNA techniques. So far, attempts to synthesize factor VIII by these techniques have been hampered by the fact that a source of factor VIII mRNA, necessary for gene cloning, is not available. If the cells described here indeed synthesize factor VIII, they should be considered as a potential source of factor VIII mRNA.

We thank Professor P. M. Mannucci for providing facilities

in his laboratory to study liver biopsy material from a patient with severe haemophilia A; Dr K. S. Sakariassen for performing the perfusion studies; Dr R. M. Bertina for the human antiserum against FVIII/C; and Professor W. G. van Aken for critically reading the manuscript. We acknowledge the technical assistance of Miss E. Cristen and the photography of Mr T. van Os.

Received 27 January; accepted 6 April 1983.

1. Zimmerman, T. S. & Meyer, D. in *Haemostasis and Thrombosis* (eds Bloom, A. L. & Thomas, D. P.) 111–123 (Churchill Livingstone, Edinburgh, 1981).
2. Hoyer, L. W. *Blood* **58**, 1–13 (1981).
3. Bloom, A. L., Giddings, J. C. & Wilks, C. J. *Nature new Biol.* **241**, 217–219 (1973).
4. Hoyer, L. W., de los Santos, R. P. & Hoyer, J. R. *J. clin. Invest.* **52**, 2737–2744 (1973).
5. Jaffe, E. A., Hoyer, L. W. & Nachman, R. L. *J. clin. Invest.* **52**, 2757–2764 (1973).
6. Nachman, R., Levine, R. & Jaffe, E. A. *J. clin. Invest.* **60**, 914–921 (1977).
7. Webster, W. P., Dodds, W. J., Mandel, S. R. & Penick, G. D. in *Handbook of Hemophilia* (eds Brinkhous, K. M. & Hemker, H. C.) 149–163 (Excerpta Medica, Amsterdam, 1975).
8. van Mourik, J. A. & Mochtar, I. A. *Biochim. biophys. Acta* **221**, 677–679 (1970).
9. Galfre, G., Howe, S. C., Milstein, C., Butcher, G. W. & Howard, J. C., *Nature* **266**, 550–552 (1977).
10. Sakariassen, K. S., Bolhuis, P. A. & Sixma, J. J. *Thromb. Res.* **19**, 547–559 (1980).
11. Tuddenham, E. G. D., Trabold, N. C., Collins, J. A. & Hoyer, L. W. *J. Lab. clin. Med.* **93**, 40–53 (1979).
12. Lansdorp, P. M., Astaldi, G. C. B., Oosterhof, F., Janssen, M. C. & Zeijlemaker, W. P. *J. immun. Meth.* **39**, 393–405 (1980).
13. Shaw, E., Giddings, J. C., Peake, I. R. & Bloom, A. L. *Br. J. Haemat.* **41**, 585–596 (1979).
14. Hellings, J. A., van Leeuwen, F. R., Over, J. & van Mourik, J. A. *Thromb. Res.* **26**, 297–302 (1982).
15. Naot, Y. & Remington, J. S. *J. immun. Meth.* **43**, 333–341 (1981).
16. Muller, H. P., van Tilburg, N. H., Bertina, R. M., Terwiel, J. Ph. & Veltkamp J. J. *Clinica chim. Acta* **107**, 11–19 (1980).
17. Muller, H. P., van Tilburg, N. H., Derks, J., Klein, E. & Bertina, R. M. *Blood* **58**, 1000–1006 (1981).
18. Solà, B., Avner, P., Sultan, Y., Jeanneau, C. & Maisonneuve, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 183–187 (1982).

Characterization of rat hypothalamic growth hormone-releasing factor

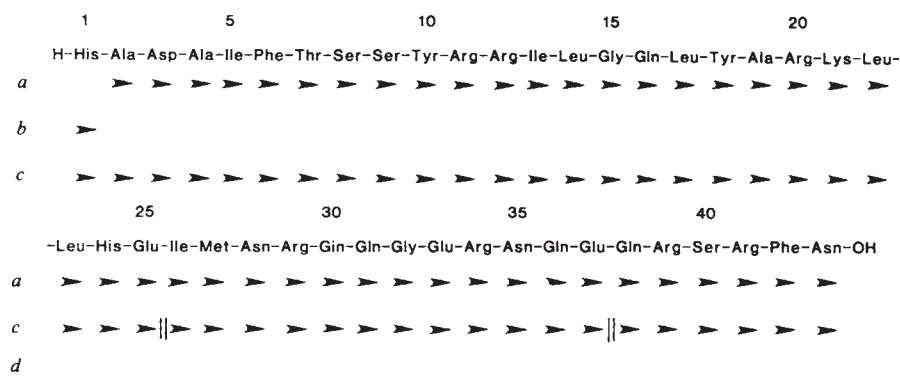
Joachim Spiess, Jean Rivier & Wylie Vale

Peptide Biology Laboratory, The Salk Institute, PO Box 85800, San Diego, California 92138, USA

The importance of the hypothalamus in the regulation of growth hormone (GH) secretion from the adenohypophysis has been supported by experimental and clinical evidence^{1–3}. It has been observed that GH secretion is in part controlled by hypothalamic GH release-inhibiting factors⁴, two forms of which, somatostatin-14 (ref. 5) and somatostatin-28 (refs 6–9), have been characterized. Peptides with GH-releasing activity were recently purified from two human pancreatic tumours of acromegalic patients and characterized^{10–14}. Synthetic versions of these human pancreatic tumour GH-releasing factors (hpGRFs) were found to be potent GH secretagogues *in vitro*^{11,12,15} and *in vivo*^{12,16,17} and to exhibit actions and interactions similar to those observed with partially purified hypothalamic GRF¹². GH-releasing peptides have been purified from hypothalami of various species^{4,15,18,36} but the structure of natural hypothalamic GRF has not yet been reported¹⁹. We now describe the purification, sequence analysis and synthesis of a 43-residue polypeptide isolated from rat hypothalamic extracts on the basis of its ability to stimulate GH secretion from cultured rat adenohypophyseal cells. The native polypeptide and its synthetic replicate do not differ significantly in their biological potencies and intrinsic activities. We propose that this polypeptide, which shows 67% homology with the corresponding N-terminal 43 residues in hpGRF-44, is the long-sought rat hypothalamic GRF (rhGRF).

The purification of GRF from rat hypothalamic tissue provided by Dr A. Parlow was monitored for GH-releasing activity with the rat anterior pituitary cell culture assay²⁰. The addition of glucocorticoids or triiodothyronine to the cultured adenohypophyseal cells enhanced the sensitivity to GRF and decreased the sensitivity to somatostatin¹⁵. Several batches totalling approximately 80,000 hypothalami were defatted with acetone,

Fig. 1 Primary structure of rhGRF and strategy of sequence analysis. ►, Identified residue or functional group. *a*, 0.9 nmol rhGRF (30.0 μ M) was incubated with 40 mM 3-sulphophenylisothiocyanate (3-SPITC) as described previously¹³ and subjected to Edman degradation in a Wittmann-Liebold³⁰ modified spinning cup sequencer^{21,22} containing Polybrene as peptide carrier. The phenylthiohydantoin (PTH) amino acids were determined with reverse-phase HPLC (minimal detectable amount <10 pmol, coefficient of variation 1–6%). PTH-amino acid yields ranged from 400 (residue 5) to 32 (residue 43) pmol. *b*, 200 pmol of GRF were hydrolysed in constant boiling HCl containing norleucine as internal standard and 3 μ l ml⁻¹ β -mercaptoethanol (3 h, 140 °C). The hydrolysate was analysed for amino acids²¹. The following amino acid ratios (nearest integer in parentheses) were determined: Asx, 3.9 (4); Thr, 1.0 (1); Ser, 3.0 (corrected for losses during hydrolysis) (3); Glx, 8.0 (8); Gly, 2.1 (2); Ala, 2.8 (3); Met, 0.8 (1); Ile, 2.8 (3); Leu, 4.1 (4); Tyr, 1.9 (2); Phe, 1.9 (2); Lys, 1.2 (1); His, 2.0 (2); Arg, 6.6 (7). Compared with the data described in *a*, one additional histidine was identified. *c*, The complete amino acid sequence was confirmed by peptide mapping (Fig. 2). *d*, The free C-terminal COOH-group was established by chromatographic comparison of the C-terminal fragment rhGRF(38–43)-OH (cleaved from the native peptide with staphylococcal protease V8) with synthetic rhGRF(38–43)-NH₂, rhGRF(38–43)-OH and C-terminally extended analogues on reverse-phase HPLC (Fig. 2).



extracted in acid-reducing conditions and gel-filtered through a Sephadex G-50 column as described earlier¹⁵. The addition of β -mercaptoethanol as reducing agent was crucial to conserve GH-releasing activity and was therefore continued throughout the whole purification. Fractions eluted from the G-50 column with partition coefficients of K_{av} = 0.20–0.31 contained biological GH-releasing activity and biological and immunological corticotropin-releasing activity. The zone comprising these fractions was separated from zones containing somatostatin-28 and somatostatin-14. The gel-filtered rhGRF (10,000–20,000 hypothalamic equivalents in 4 l of filtrate) was purified batchwise on a preparative reverse-phase HPLC cartridge (5 $\times$ 30 cm) in a Waters Associates Prep 500 system. The cartridge was packed with octadecyl-bonded silica of particle size 20 μ m and eluted at a flow rate of 75 ml min⁻¹ with a mixture of solvents A and B (7–60% B in 40 min). Solvent A was 0.08 M phosphoric acid adjusted with triethylamine to pH 2.25 (TEAP); solvent B consisted of solvent A (40%) and acetonitrile (60%). rhGRF was separated from CRF in this experiment. Final purification was accomplished with semi-preparative and analytical reverse-phase HPLC, which was performed in a Beckman model 100A liquid chromatograph equipped with a Kratos Spectroflow 773 detector and Spectraphysics Minigrator. The samples were applied to end-capped Vydac (0.46 $\times$ 25 cm) columns containing octadecyl (C₁₈) or diphenyl groups bonded to spherical silica of particle size 5 μ m and pore size 330 Å. The columns were eluted at 23 °C and a flow rate of 1.2 ml min⁻¹ with mixtures of solvents A and B, as described above, or with 0.1% (v/v) trifluoroacetic acid instead of TEAP. Based on the bioassay analysis of all purification steps, only one predominant form of rhGRF was observed. However, it cannot be excluded that the hypothalamic extract contained minor amounts of other forms of GRF. A total of approximately 3.4 nmol of rhGRF, as determined by amino acid analysis, was obtained. Thus, approximately 43 fmol of rhGRF were purified from one rat hypothalamus.

An aliquot of the purified peptide (0.9 nmol) was subjected to Edman degradation in an automatic spinning cup sequencer^{21,22}. All 43 residues of rhGRF, with the exception of the N-terminal residue, were determined in this experiment (Fig. 1). The results were confirmed and supplemented with data of peptide mapping using reverse-phase HPLC of 1.4 nmol of rhGRF digested with staphylococcal protease (Fig. 2). The proteolytic fragments were identified by chromatographic comparison with the corresponding synthetic replicates, amino acid analysis and Edman degradation. The sequence (Fig. 1) was further validated by the observation that the amino acid analysis of the rhGRF hydrolysate was in agreement with the amino acid composition of rhGRF as determined by sequence analysis (see Fig. 1 legend).

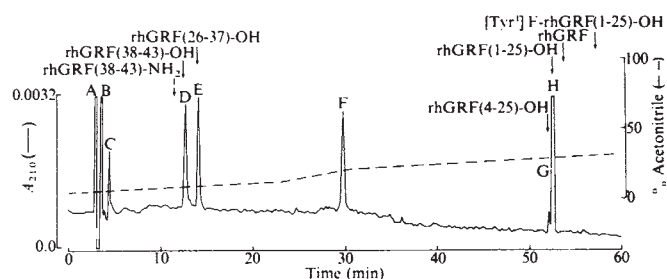


Fig. 2 Peptide mapping of GRF with reverse-phase HPLC carried out with a Hewlett-Packard 1084B liquid chromatograph. rhGRF (1.4 nmol) was digested with staphylococcal protease V8 (Miles) in 50 μ l of 50 mM *N*-ethylmorpholine adjusted to pH 7.8 with trifluoroacetic acid (TFA; substrate/enzyme weight ratio 20:1; 2 h at 37 °C). The enzyme cleaves mainly at the C-terminal side of glutamic acid and to a lesser extent aspartic acid residues³¹. The digest was applied to a Vydac C₁₈ column (0.46 $\times$ 25 cm; particle size 5 μ m; pore size 330 Å) and eluted at 30 °C and a flow rate of 1.0 ml min⁻¹ with a mixture of 0.05% TFA and acetonitrile. Based on absorbance of the eluate at 210 nm, products A–H were detected. Protein, as determined by amino acid analysis, was only found in fractions D (0.7 μ g), E (1.1 μ g), G (<0.1 μ g) and H (1.8 μ g) in agreement with the data of the fingerprint of synthetic rhGRF performed in the same conditions. The three main peptide products D, E and H were identified as rhGRF(38–43)-OH, rhGRF(26–37)-OH and rhGRF(1–25)-OH respectively, by retention times, amino acid analysis and Edman degradation (complete analysis of 0.4–0.7 nmol of peptide in the spinning cup). Retention times of synthetic peptides or their fragments are marked with arrows. The appearance of rhGRF(26–37)-OH (fraction E) containing glutamic acid-33 is consistent with our finding that the peptide bond between Glu 33 and Arg 34 of synthetic rhGRF and its analogues was resistant to cleavage by staphylococcal protease in the described conditions.

That the characterized 43-residue polypeptide was responsible for the assayed biological activity was supported by the finding that native rhGRF and its synthetic replicate (synthesis described in Fig. 3 legend) did not differ significantly in their biological activities *in vitro* (Fig. 3). Native and synthetic rhGRF were highly potent stimulators of the GH secretion *in vitro* with EC₅₀ values in the range 10–20 pM and minimal detectable concentrations of <4 pM. rhGRF appeared to be slightly more potent than hpGRF-44 in the rat anterior pituitary cell culture assay. As observed with purified native rhGRF and hpGRF^{11,12,15,23}, the GH secretory response to synthetic rhGRF was found to be noncompetitively inhibited by somatostatin-14 and to be Ca²⁺-independent.

rhGRF belongs structurally to the glucagon–secretin family, a group of pancreatic intestinal peptides, as does hpGRF. Although most peptides of this family differ in size, comparison

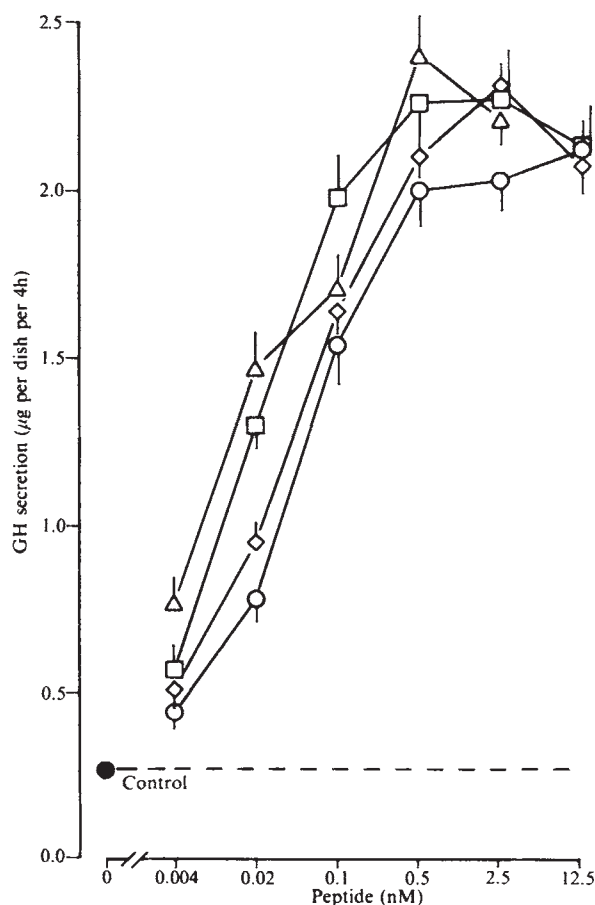


Fig. 3 Effects (potencies, with 95% confidence limits, in parentheses) of: Δ , native rhGRF(3.38:2.07–5.78); $\square$, synthetic rhGRF(2.67:1.80–4.07); $\circ$, synthetic hpGRF-40(1.0); $\diamond$, synthetic hpGRF-44(1.38:1.02–1.87) on GH secretion from rat adenohypophyseal cells in a 4-day-old monolayer culture. Cells were dissociated and established in culture in β -PJ medium which contained 2% fetal bovine serum, 5 nM dexamethasone and 30 pM triiodothyronine¹⁵. Serum-free medium containing 0.1% bovine serum albumin and 0.5 mM ascorbic acid was used during the 4-h incubation. GH levels in the incubation medium were measured with radioimmunoassay kits distributed by The National Hormone and Pituitary Program of NIADDK. Standard deviations of the mean are indicated. Peptides were synthesized with the solid-phase approach using experimental procedures similar to those described earlier^{32–34}. Protecting groups (Tos, tosyl; Bzl, benzyl; 2,6Cl₂Bzl, 2,6 dichlorobenzyl; Xan, xanthryl; 2ClZ, 2-chlorobenzylloxycarbonyl) for N-terminal or side-chain protection were those shown below for derivatized rhGRF: Boc-His(Tos) - Ala - Asp(Bzl) - Ala - Ile - Phe - Thr(Bzl) - Ser(Bzl) - Ser(Bzl) - Tyr(2,6Cl₂Bzl) - Arg(Tos) - Arg(Tos) - Ile - Leu - Gly - Gln(Xan) - Leu - Tyr(2,6Cl₂Bzl) - Ala - Arg(Tos) - Lys(2ClZ) - Leu - Leu - His(Tos) - Glu(Bzl) - Ile - Met - Asn(Xan) - Arg(Tos) - Gln(Xan) - Gln(Xan) - Gly - Glu(Bzl) - Arg(Tos) - Asn(Xan) - Gln(Xan) - Glu(Bzl) - Gln(Xan) - Arg(Tos) - Ser(Bzl) - Arg(Tos) - Phe - Asn(Xan)O - CH₂ - [resin]. After complete deprotection and cleavage with HF, the crude preparation of rhGRF was purified on analytical reverse-phase HPLC columns in similar conditions as described for the isolation of the native peptide. Based on chromatographic behaviour in reverse-phase HPLC, amino acid analysis and peptide mapping (as described for native rhGRF), the purity of the synthetic product is estimated to be >95%. Synthetic rhGRF co-eluted in reverse-phase HPLC with the native peptide.



Fig. 4 Sequence comparison of rhGRF with hpGRF-44 and other members of the glucagon-secretin family, isolated from hog (PHI, peptide with histidine as N-terminus and isoleucine amide as C-terminus; VIP, vasoactive intestinal peptide; GIP, gastric inhibitory peptide). Homologies between rhGRF and other peptides are shaded. ■, C-terminal amide; A, alanine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

of the sequences of the smaller peptides with corresponding N-terminal sequences of the larger peptides shows significant homologies. Thus, the sequence homologies of rhGRF with hpGRF-40, hpGRF-44 and porcine PHI (PHI-27)²⁴, vasoactive intestinal peptide²⁵, glucagon²⁶, secretin²⁸ and gastric inhibitory peptide (GIP)²⁷ were calculated to be 70, 67, 52, 43, 31, 30 and 14% respectively (Fig. 4). None of the members of this family, except rhGRF and hpGRF, exhibits detectable GH-releasing activity as described earlier^{11,12}. The 14 amino acid differences between rhGRF and hpGRF-44 are distributed over the whole molecule, but especially accumulated in the C-terminal end. All differences except two (Phe 42 versus Ala 42 and Asn 43 versus Arg 43) could be explained as the result of single-base mutations. As hpGRF-44 was reported¹¹ to be C-terminally amidated, we do not exclude the possibility (which formally always has to be considered for peptides with a free C-terminus) that the characterized rhGRF was enzymatically cleaved to the 43-residue peptide during extraction or purification. In view of the C-terminal heterogeneity of GRF-like peptides isolated from two tumours^{11,12} and the high potency of shortened forms of hpGRF^{11,12} and rhGRF, such as rhGRF(1–29)-NH₂ (J. R., J. S. and W. V., in preparation), it is conceivable that several forms of hypothalamic GRF might be produced, as has been shown to be the case for somatostatin. Whether the significant sequence difference between human

pancreatic tumour GRF and rhGRF reflects tissue or species variation has to be clarified by sequence determination of human hypothalamic GRF. Immunocytochemical findings with antisera raised against hpGRF-40 show the presence of peptides related to hpGRF in fibres and cell bodies of the primate³⁵ and rat hypothalamus (L. Swanson *et al.*, in preparation). We have recently isolated two GH-releasing peptides from the human stalk median eminence (provided by Dr S. Reichlin). The two peptides, oxidized to the corresponding methionine sulfoxide form (to facilitate their purification), appeared to be identical to hpGRF-44 and hpGRF-40, respectively, on the basis of the reverse-phase HPLC behaviour of the intact peptides and the clostripain digest (peptide mapping). Proof of identity has to be provided by complete sequence analysis. It is interesting that rhGRF is more closely related to the other members of the glucagon-secretin family, with the exception of GIP, than is hpGRF. Accordingly, five peptides (including rhGRF) of this family carry histidine as the N-terminus and only GIP and hpGRF contain tyrosine as the N-terminal residue.

Based on the described chemical and biological data, we present here for the first time the primary structure of a GRF established to be of hypothalamic origin and thus representing the long-sought hypothalamic growth hormone-releasing factor, which probably plays a major part in the regulation of GH secretion in agreement with Harris²⁹ hypothalamic portal

vessel-chemotransmitter hypothesis.

This research was supported in part by NIH grants AM-26741, AM20971, HD13527, and by grants from The Rockefeller Foundation and The March of Dimes-Birth Defects Foundation. Sequence analysis facilities were supported in part by the Stiftung Volkswagenwerk and Fritz Thyssen Stiftung. Research was conducted in part by The

Clayton Foundation for Research, California Division. J. S. and W. V. are Clayton Foundation Investigators. We thank Dr A. Parlow and The National Hormone and Pituitary Program and Dr S. Reichlin for providing hypothalamic tissue, R. Lee, J. Heil, J. Vaughan, R. Galyean, G. Yamamoto, R. Kaiser and R. McClintock for technical assistance and L. Wheatley, D. Weigand and S. McCall for manuscript preparation.

Received 28 March; accepted 28 April 1983.

1. Reichlin, S. *Endocrinology* **69**, 225-230 (1961).
2. Martin, J. B. *et al.* in *The Hypothalamus* (eds Reichlin, S., Baldessarini, R. & Martin, J. B.) 329-355 (Raven, New York, 1978).
3. Frohman, L. A. & Szabo, M. in *Physiopathology of Endocrine Diseases and Mechanisms of Hormone Action* (eds Soto, R. J., Blaquier, J. & deNicola, A.) 259-271 (Liss, New York, 1981).
4. Krulich, L., Dhariwal, A. P. S. & McCann, S. M. *Endocrinology* **83**, 783-790 (1968).
5. Brazeau, P. *et al.* *Science* **179**, 77-79 (1973).
6. Pradayrol, L., Jörnvall, H., Mutt, V. & Ribet, A. *FEBS Lett.* **109**, 55-58 (1980).
7. Schally, A. V. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 4489-4493 (1980).
8. Esch, F., Böhlen, P., Ling, N., Benoit, R. & Guillemin, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6827-6831 (1980).
9. Spiess, J., Villarreal, J. & Vale, W. *Biochemistry* **20**, 1981-1988 (1981).
10. Thorner, M. O. *et al.* *J. clin. Invest.* **70**, 965-977 (1982).
11. Guillemin, R. *et al.* *Science* **218**, 585-587 (1982).
12. Rivier, J., Spiess, J., Thorner, M. & Vale, W. *Nature* **300**, 276-278 (1982).
13. Spiess, J., Rivier, J., Thorner, M. & Vale, W. *Biochemistry* **24**, 6037-6040 (1982).
14. Esch, F. S. *et al.* *Biochem. biophys. Res. Commun.* **109**, 152-158 (1982).
15. Vale, W., Vaughan, J., Yamamoto, G., Spiess, J. & Rivier, J. *Endocrinology* **112**, 1553-1555 (1983).
16. Thorner, M. O. *et al.* *Lancet* **i**, 24-28 (1983).
17. Wehrenberg, W. B. *et al.* *Biochem. biophys. Res. Commun.* **109**, 382-387 (1982).

18. Sykes, J. E. C. & Lowry, P. J. *Biochem. J.* **209**, 643-651 (1983).
19. Buckingham, J. C. *Prog. med. Chem.* **15**, 166-210 (1978).
20. Vale, W. *et al.* *Recent Prog. Horm. Res.* **31**, 365-397 (1975).
21. Spiess, J., Rivier, J., Rivier, C. & Vale, W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6517-6521 (1981).
22. Spiess, J., Rivier, J., Rivier, C. & Vale, W. in *Methods in Protein Sequence Analysis* (ed. Elzinga, M.) 131-138 (Humana, Clifton, New Jersey, 1982).
23. Brazeau, P. *et al.* *Biochem. biophys. Res. Commun.* **109**, 588-594 (1982).
24. Tatemoto, K. & Mutt, V. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6603-6607 (1981).
25. Mutt, V. & Said, S. I. *Eur. J. Biochem.* **42**, 581-589 (1974).
26. Bromer, W. W., Sinn, L. G. & Behrens, O. K. *J. Am. chem. Soc.* **79**, 2807-2810 (1957).
27. Brown, J. C. & Dryburgh, J. R. *Can. J. Biochem.* **49**, 867-872 (1971).
28. Mutt, V., Jorpes, J. E. & Magnusson, S. D. *Eur. J. Biochem.* **15**, 513-519 (1970).
29. Green, J. D. & Harris, G. W. *J. Physiol., Lond.* **108**, 359 (1949).
30. Wittmann-Liebold, B. in *Polypeptide Hormones* (eds Beers, R. F. Jr & Bassett, E. G.) 87-120 (Raven, New York, 1980).
31. Houmard, J. & Drapeau, G. R. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3506-3509 (1972).
32. Merrifield, R. B. *J. Am. chem. Soc.* **85**, 2149-2156 (1963).
33. Rivier, J. *J. Am. chem. Soc.* **96**, 2986-2992 (1974).
34. Märki, W., Spiess, J., Taché, Y., Brown, M. & Rivier, J. E. *J. Am. chem. Soc.* **103**, 3178-3185 (1981).
35. Bloch, B. *et al.* *Nature* **301**, 607-608 (1983).
36. Luben, R. A., Brazeau, P., Böhlen, P. & Guillemin, R. *Science* **218**, 887-889 (1982).

Novel dihydropyridines with positive inotropic action through activation of Ca^{2+} channels

M. Schramm, G. Thomas, R. Towart & G. Franckowiak

Pharma Research Centre, Bayer AG, D-5600 Wuppertal 1, FRG

Transmembrane influx of extracellular calcium through specific calcium channels is now accepted to have an important role in the excitation-contraction coupling of cardiac and smooth muscle¹⁻³. The importance of such slow calcium channels has been underlined by the development of specific calcium channel blocking agents, the 'calcium antagonists', typified by verapamil, nifedipine and diltiazem⁴. These drugs have been used to investigate the properties of slow calcium channels in a variety of tissues. We have found that small modifications to the nifedipine molecule produce other dihydropyridine derivatives (see Fig. 1) with effects diametrically opposite to those of the calcium antagonists: cardiac contractility is stimulated and smooth muscle is contracted. These effects are competitively antagonized by nifedipine. Apparently, nifedipine and the novel compounds bind to the same specific dihydropyridine binding sites in or near the calcium channel. In contrast to nifedipine, however, the new compounds promote—instead of inhibiting—the influx of Ca^{2+} ions. We report here the properties of BAY K 8644 (methyl 1,4-dihydro-2,6-dimethyl-3-nitro-4-(2-trifluoromethylphenyl)-pyridine-5-carboxylate), one of the most potent of these novel compounds.

Nifedipine (Fig. 1) and its ortho- CF_3 analogue⁵ are potent calcium antagonists and thus possess negative inotropic and smooth muscle relaxant properties⁶⁻⁸. BAY K 8644 (Fig. 1), another nifedipine analogue⁹, in a similar dose range produces marked positive inotropic and vasoconstrictor effects, as demonstrated in the pentobarbitone-anaesthetized dog (Fig. 2); nifedipine is able to reverse these effects. We therefore examined the effects of BAY K 8644 on isolated perfused guinea pig hearts and blood vessel strips, and examined its interactions with nifedipine. In the isolated guinea pig heart, BAY K 8644 increases contractility as shown by the increase in left ventricular pressure (solid circles in Fig. 3). The positive inotropic effect begins at $10^{-9} \text{ mol l}^{-1}$ and reaches a plateau at $10^{-7} \text{ mol l}^{-1}$, that is, the same concentration range in which

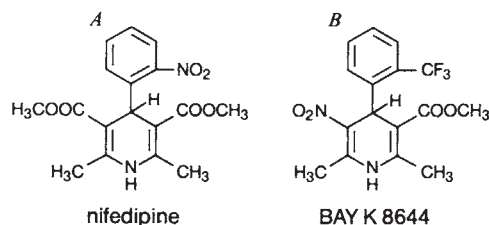


Fig. 1 Chemical structure of nifedipine (A) and BAY K 8644 (B).

nifedipine exerts its negative inotropic effect. Adding nifedipine to the perfusion medium produces a parallel shift to the right of the BAY K 8644 dose-response curve (Fig. 3). The positive inotropic effects are not due to α - or β -adrenoceptor stimulation, as neither phentolamine nor propranolol were able to inhibit these actions.

BAY K 8644 had little effect on the isolated rabbit aortic strip incubated in Tyrode's solution (K^+ concentration 2.6 mmol l^{-1}). However, this vessel is quiescent, and the potential-sensitive calcium channels in this tissue are probably not activated in resting conditions². When the vessels are partly depolarized by adding $15 \text{ mmol l}^{-1} \text{ K}^+$ (still sub-threshold for K^+ -induced contractions), BAY K 8644 produces concentration-dependent contractions in concentrations comparable to those observed for positive inotropic effects in the isolated heart (Fig. 4). Figure 4 also shows the effects of BAY K 8644 in the presence of increasing nifedipine concentrations. Nifedipine concentration dependently inhibits the effects of BAY K 8644, producing a parallel shift of the dose-response curve to the right, indicative of competitive antagonism. In contrast, verapamil and diltiazem, two calcium antagonists not of the dihydropyridine-type, also inhibit the BAY K 8644-induced effects, but not competitively, as the maxima of the dose-response curve to BAY K 8644 were markedly and dose-dependently depressed in the presence of the two drugs (R.T. *et al.* in preparation).

In an attempt to examine the effect of BAY K 8644 on calcium movements, isolated rabbit aortic strips were incubated in low calcium solution (0.06 mmol l^{-1}), depolarized by addition of K^+ (40 mmol l^{-1}), and the relation between contraction and added calcium was investigated. The results plotted in a Lineweaver-Burk diagram^{10,11} indicate that nifedipine simulates a competitive inhibition of the calcium movements through calcium channels, and that BAY K 8644 enhances calcium influx by interacting with the same site. On the other hand,

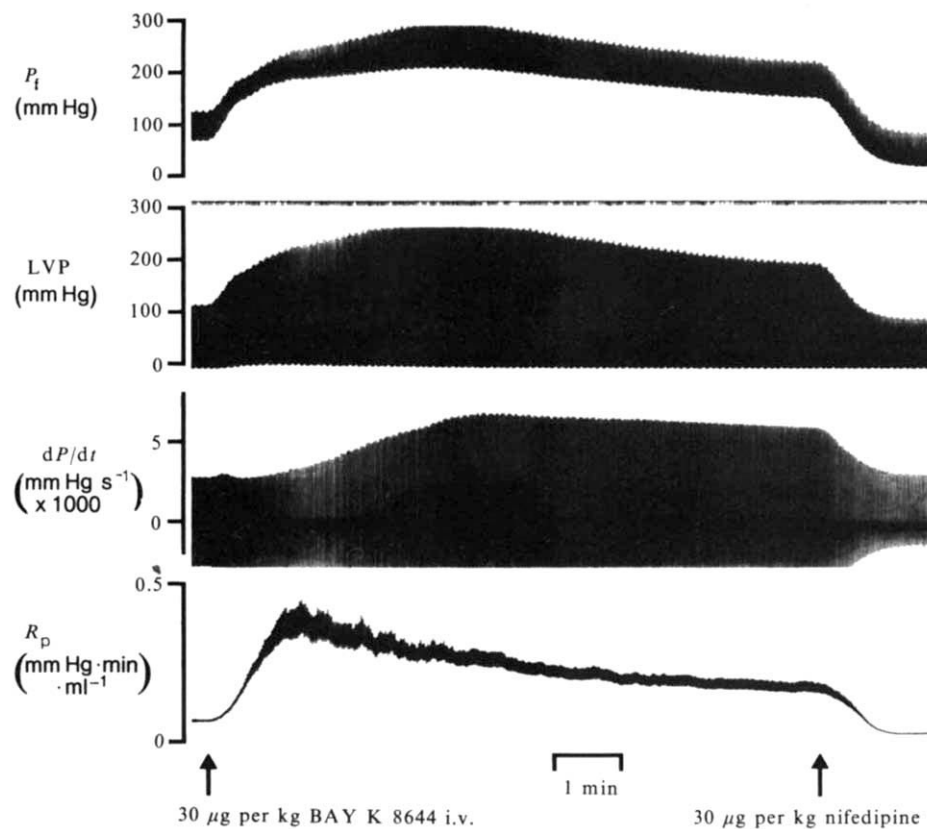


Fig. 2 Effects of a high dose of BAY K 8644 on blood pressure in the femoral artery (P_f), pressure in the left ventricle (LVP) and its first derivative (dP/dt) as a parameter for the contractility of the heart (negative edge clipped), and the peripheral resistance (R_p , calculated as $\dot{V}_{\text{aorta}}/P_f$) of the pentobarbitone-anaesthetized mongrel dog (for experimental details see ref. 16). These effects are reversed by a similar dose of nifedipine; i.v., intravenous.

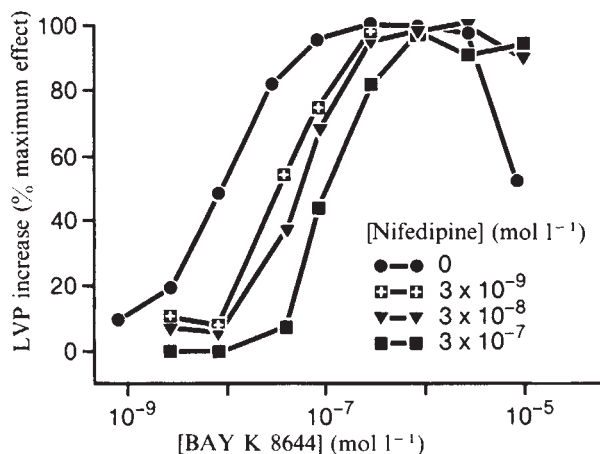


Fig. 3 Dose-response curves (for isolated perfused guinea pig heart, left ventricular pressure) to BAY K 8644 for different nifedipine concentrations. Hearts of adult animals (250–300 g) were perfused according to the Langendorff method as modified by Ope¹⁷ at 32°C with Krebs–Henseleit solution at a constant flow rate (10 ml min^{-1}). The contractions of the left ventricle were recorded isovolumetrically (left ventricular balloon pressure). Drug solutions were infused into the aortic cannula at 0.1 ml min^{-1} . The maximum effect of BAY K 8644 was taken as 100%. Each point is the mean value of four observations.

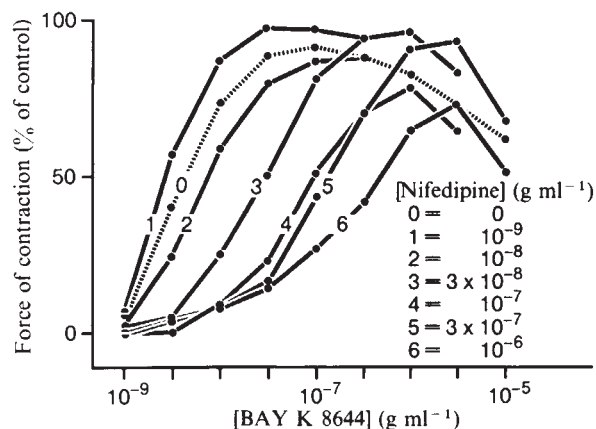


Fig. 4 Contractions of the rabbit aortic strip for increasing BAY K 8644 concentrations. Strips of the rabbit aorta¹⁸ were suspended in Tyrode's solution (37°C) at sub-threshold K^+ concentration (15 mmol l^{-1}). The Bay K 8644-induced contractions were recorded using strain gauges, and are expressed as a percentage of previous control contractions induced by $50 \text{ mmol l}^{-1} \text{K}^+$ (see ref. 19 for details). Each point is the mean of at least three observations.

BAY K 8644 does not bind specifically to other known receptors such as the dopamine, muscarinic, α - and β -adrenergic, serotonin, histamine, γ -aminobutyric acid, diazepam and opiate receptors (P. Bellemann, unpublished observations). Recently, the existence of a specific dihydropyridine receptor has been demonstrated using ^3H -labelled dihydropyridines by different laboratories^{12–15,20}.

As BAY K 8644 also constricts the coronary arteries (M.S. *et al.*, in preparation), this compound does not seem useful as a therapeutic agent or as an antidote for calcium antagonists.

The observations reported here provide the first evidence that there exist some sites in or near the calcium channel which

bind dihydropyridine derivatives, and that occupation of these sites may either increase ('calcium agonists', for example, BAY K 8644) or decrease ('calcium antagonists', for example, nifedipine) the transmembrane calcium flux into the cell. These binding sites are specific for dihydropyridines as calcium antagonists of different chemical structure such as verapamil or diltiazem do not interfere competitively with the effect of BAY K 8644.

The potency and selectivity of such exogenous calcium channel modulators suggest that there may be some endogenous factor which operates or modulates physiologically calcium channel function.

Received 2 March; accepted 30 March 1983.

1. Kohlhardt, M. *Basic Res. Cardiol.* **76**, 589–601 (1981).
2. Bolton, T. B. *Physiol. Rev.* **59**, 607–718 (1979).
3. Reuter, H. *Nature* **301**, 569–574 (1983).
4. Fleckenstein, A. A. *Rev. Pharmac. Tox.* **17**, 149–166 (1977).
5. Loev, B., Ehrreich, S. J. & Tedeschi, R. E. *J. Pharm. Pharmac.* **24**, 917–918 (1972).
6. Loev, B., Goodman, M. M., Snader, K. M., Tedeschi, R. & Macko, E. *J. med. Chem.* **17**, 956–965 (1974).
7. Rodenkirchen, R. et al. *Naunyn-Schmiedeberg's Archs Pharmac.* **310**, 69–78 (1979).
8. Triggie, D. J. & Swany, V. C. *Circulation Res.* (in the press).
9. Germ. Pat. Applic. DOS 3130041 (1981).
10. Lineweaver, H. & Burke, D. *J. Am. chem. Soc.* **56**, 658–666 (1934).
11. Rosenberger, L. B., Ticku, M. K. & Triggie, D. J. *Can. J. Physiol. Pharmac.* **57**, 333–347 (1979).
12. Bellemann, P., Ferry, D., Lübbecke, F. & Glossmann, H. *Arzneimittel-Forsch.* **31**, 2064–2067 (1981).
13. Bolger, G. T. et al. *Biochem. biophys. Res. Commun.* **104**, 1604–1609 (1982).
14. Murphy, K. M. M. & Snyder, S. H. *Eur. J. Pharmac.* **77**, 201–202 (1982).
15. Holck, M., Thorens, S. & Haeusler, G. *Eur. J. Pharmac.* **85**, 305–315 (1982).
16. Vater, W. et al. *Arzneimittel-Forsch.* **22**, 1–14 (1972).
17. Opie, L. H. *J. Physiol., Lond.* **180**, 529–541 (1965).
18. Furchgott, R. F. & Bhadrakom, S. *J. Pharmac. exp. Ther.* **108**, 129–143 (1953).
19. Towart, R. *J. Cardiovasc. Pharmac.* **4**, 895–902 (1982).
20. Bellemann, P., Schade, A. & Towart, R. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2356–2360 (1983).

Bipolar cells in the mudpuppy retina use an excitatory amino acid neurotransmitter

Malcolm M. Slaughter & Robert F. Miller

Departments of Ophthalmology, Physiology, and Biophysics,
Washington University School of Medicine, St Louis,
Missouri 63110, USA

The bipolar cells of the vertebrate retina are the principal neuronal elements which transmit photoreceptor activity from the outer to the inner retina¹. An important function of the bipolars is to segregate photoreceptor input into independent ON and OFF channels^{2,3} which are subserved, respectively, by the depolarizing and hyperpolarizing bipolar subtypes^{4–7}. Ultrastructural^{8,9} and physiological¹⁰ observations suggest that chemical neurotransmission is the predominant means of bipolar input to the inner retina. Both ON and OFF bipolars apparently release excitatory transmitters^{11–13}. Histological studies with cytotoxic agents^{14–16} and physiological studies^{17,18} indicate that third-order neurones have excitatory amino acid receptors. In ON–OFF amacrine and ganglion cells, which receive input from both bipolars^{12,19}, ON and OFF excitation have a similar ionic basis²⁰, suggesting that the same transmitter may be released by both types of bipolars. We have now found that (±)c/s-2,3-piperidine dicarboxylic acid (PDA)²¹, a new excitatory amino acid antagonist, blocks bipolar input to the inner retina and thus suggests that an excitatory amino acid is a bipolar cell transmitter.

Experiments were conducted in the superfused retina eyecup of the mudpuppy, *Necturus maculosus*, using standard intracellular recording techniques and diffuse, focal (200 µm) or annular (i.d. = 400 µm, o.d. > retinal diameter) light stimulation with an irradiance of 5×10^{-9} W cm⁻² (see ref. 12 for details of the methodology). Cells were identified from depth of electrode penetration and light response characteristics.

We recently reported that PDA blocked the light responses of horizontal cells and OFF bipolars but not ON bipolars²². In the proximal retina, this would result in a selective depression of the OFF responses while ON activity should persist. However, as seen in Fig. 1, intracellular recordings from different types of amacrine and ganglion cells show that PDA application also suppresses ON responses. This indicates that PDA had a direct effect in the inner retina.

Figure 1a shows the light response of a sustained ON ganglion cell. This neurone receives excitatory input from ON bipolars¹⁹, probably through direct synaptic connections^{23,24}. The application of 6 mM PDA caused a large suppression in the amplitude of the synaptic excitatory postsynaptic potential (e.p.s.p.) and an associated reduction in spike activity. Both the response amplitude and the spike activity quickly recover after removal of PDA. We analysed the conductance changes associated with the action of PDA to distinguish between a block of excitatory synaptic input and a suppression of the light response through the activation of a shunting mechanism. The latter would be accompanied by a large (~100 MΩ) decrease in resistance²⁵. Results from a typical experiment are shown in Fig. 1b, which illustrates representative voltage responses and resistance measurements in another sustained ON ganglion cell evaluated before, during and after the application of 6 mM PDA. Although PDA caused a large reduction in the light-evoked e.p.s.p., there was no change in the voltage deflection produced by a 0.1 nA negative current pulse applied through the recording electrode. After recovery of the light response on return to control, the input resistance of the ganglion cell remained unchanged. Thus, the PDA-induced suppression of the e.p.s.p. was not the result of low resistance shunting, indicating that PDA blocked synaptic transmission from the ON bipolar to the ganglion cell.

The ON activity of transiently responding third-order neurones was also blocked by PDA. For example, Fig. 1c shows the effect of PDA on responses of a transient ON–OFF amacrine cell to focal and annular illumination. The application of 5 mM PDA resulted in a strong suppression of the ON responses and an almost total elimination of the OFF responses. The effect was rapidly reversible on return to control perfusate. A similar effect is seen in transient ON–OFF ganglion cells, as exemplified in Fig. 1d. In this neurone, biphasic current pulses (0.1 nA, negative and then positive) were concatenated with small spot and annular stimuli. When 4 mM PDA was applied there was a marked suppression of ON and OFF activity; again

Fig. 1 The effect of PDA on light-evoked activity of amacrine and ganglion cells in the mudpuppy retina. PDA suppressed the light response of both sustained ON (ON GC) and OFF (OFF GC) ganglion cells and of transient amacrine (ON–OFF AM) and ganglion (ON–OFF CC) cells. This blocking action was not associated with a change in the input resistance (arrows in b). The light stimulus was diffuse in a and b and marked by square pulses below each voltage trace. The light stimuli in c–e were an annulus (upward square pulse) and a small spot (downward square pulse). The duration of the drug application is indicated by the bar above each trace. Calibration: 20 mV in a and b, 30 mV in c–e.

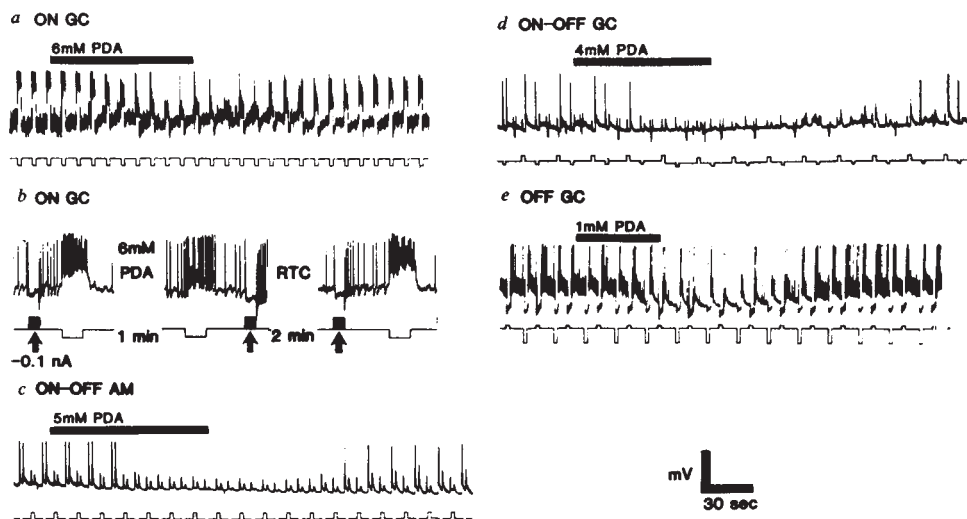
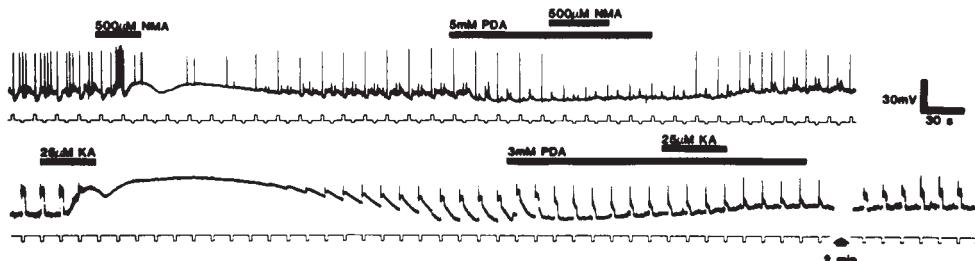


Fig. 2 The antagonistic action of PDA on the amino acid-induced depolarization of inner retinal neurones. At the left of each trace, the action of the agonist alone is shown, and at the right is the action of the same concentration of the agonist in the presence of PDA. PDA totally blocked the action of *N*-methyl-DL-aspartate (NMA) and greatly suppressed the action of kainic acid (KA). The bars above each trace indicate the duration of drug application and the square pulses below each voltage trace indicate the occurrence of light stimuli which were annuli (downward pulse) and small spots (upward pulse) in the top figure and a diffuse light in the lower trace.



the latter was more pronounced. Intracellular current pulses applied in the dark indicated no change in input resistance during the course of PDA's action on this cell. Ikeda and Sheardown¹⁷ have recently reported that aspartate mediates sustained ganglion cell activity while acetylcholine mediates transient ganglion cell activity²⁶ in the cat retina. In some species acetylcholine is an excitatory transmitter in the inner retina, but studies in the rabbit indicate that acetylcholine is released by a group of amacrine cells²⁷. Acetylcholine does not seem to have a similar role in mudpuppy. Thus, both transient and sustained activity in the mudpuppy retina can be blocked by an excitatory amino acid antagonist.

The analysis of the OFF channel input to the inner retina is less clear because PDA blocks the photoreceptor to OFF bipolar synapse²², causing a hyperpolarization of this bipolar type and thereby reducing the drive for its transmitter release. This makes it difficult to assess whether PDA can also antagonize the action of the OFF bipolar neurotransmitter. In an attempt to resolve this we used low concentrations (1 mM) of PDA which produced only a slight block of the OFF bipolar. Sustained OFF ganglion cells probably receive direct excitatory input from OFF bipolars²⁴, which is maximal in the dark. In the sustained OFF ganglion cell of Fig. 1e, 1 mM PDA caused a 75% reduction of the OFF bipolar-mediated depolarization and an elimination of the sustained spike activity. This greater effectiveness of PDA on sustained OFF ganglion cells versus OFF bipolar suggests that PDA may block the OFF bipolar neurotransmitter. However, this result could be explained by mechanisms other than effects on the OFF bipolar itself.

The conductance measurements carried out here indicate that PDA blocks synaptic transmission, but such experiments do not distinguish between a presynaptic and a postsynaptic site of action. In the outer retina, the persistence of photoreceptor to ON bipolar synaptic transmission indicated that PDA was not acting presynaptically to reduce neurotransmitter release²². To confirm this in the inner retina we evaluated the ability of PDA to block the depolarization of third-order neurones induced by exogenously applied excitatory amino acid agonists. We have found that third-order neurones in the mudpuppy retina have intrinsic kainic acid and *N*-methyl aspartate receptors²⁸. These are discrete receptor types, as D- α -amino adipate can block the *N*-methylaspartate receptor but not the kainic acid receptor. We therefore tested the action of PDA at each of these receptor sites. In the ON-OFF ganglion cell illustrated at the top of Fig. 2, 500 μ M *N*-methyl-DL-aspartate caused a prominent depolarization and a loss of the light response. After recovery, 5 mM PDA was applied, resulting in a diminution of the light responses. During continuous application of PDA, the reapplication of 500 μ M *N*-methyl-DL-aspartate had no apparent effect.

PDA also blocked the effect of kainic acid, as illustrated in the ON ganglion cell at the bottom of Fig. 2. Initially, the application of 25 μ M kainic acid produced a large and prolonged depolarization with a concomitant loss of the light response. Following recovery, 3 mM PDA was applied and caused a reduction in the sustained e.p.s.p.s, although transient ON spikes persisted. In the continued presence of PDA, 25 μ M kainic acid generated only a slight depolarization. We found

that PDA antagonized the actions of kainic acid and *N*-methylaspartate on all third-order neurones tested.

We have previously found that PDA is specific for excitatory amino acid receptors²². From these results we conclude that PDA antagonizes two distinct types of excitatory amino acid receptors and that this antagonism is the probable mode of action underlying the effects of PDA in the inner retina. PDA seems to block the action of several excitatory amino acids, thus precluding a more precise differentiation of the synaptic receptors with this agent.

As PDA blocks ON responses in all third-order neurones that we studied while the ON bipolar functions normally, we conclude that the ON bipolar neurotransmitter is an excitatory amino acid. The effectiveness of lower doses of PDA on the OFF pathway in the inner retina, coupled with the known similarity in the ON and OFF e.p.s.p.s of transient amacrine and ganglion cells, hints that the OFF bipolar may also use an excitatory amino acid neurotransmitter. Although PDA does not permit an identification of the particular amino acid involved, the inability of α -amino adipate or α -aminosuberate to block the light-evoked activity in the inner retina suggests that aspartate is not the bipolar cell transmitter in the mudpuppy retina. These results, combined with previous findings^{19,29,30}, suggest that neurotransmission at the photoreceptor synapse and at the bipolar cell synapse is subserved by glutamate or its analogue.

We thank Dr J. C. Watkins for providing the PDA used throughout most of our experiments. This chemical is now commercially available from Cambridge Research Biochemicals. We thank Rita Drochelman for typing the manuscript and Arthur Reel for technical assistance. This research was supported in part by Fight for Sight Inc., New York City and NEI grants R 23 EY 03868 and R 01 EY 03014.

Received 1 December 1982; accepted 15 April 1983.

1. Cajal, R. Y. *The Structure of the Retina* (Thomas, Springfield, 1972).
2. Werblin, F. S. & Dowling, J. E. *J. Neurophysiol.* **32**, 339-355 (1969).
3. Kaneko, A. *J. Physiol., Lond.* **207**, 623-633 (1970).
4. Miller, R. F. & Dacheux, R. F. *J. gen. Physiol.* **67**, 679-690 (1976).
5. Naka, K. I. *Invest. Ophthalmol.* **15**, 926-935 (1976).
6. Famiglietti, E. V., Kaneko, A. & Tachibana, M. *Science* **198**, 1267-1269 (1977).
7. Nelson, R., Famiglietti, E. V. & Kolb, H. *J. Neurophysiol.* **41**, 472-483 (1978).
8. Kidd, M. *J. Anat.* **96**, 179-187 (1962).
9. Dowling, J. & Boycott, B. *Proc. R. Soc. Lond.* **160**, 80-111 (1966).
10. Dacheux, R. F. & Miller, R. F. *Science* **191**, 963-964 (1976).
11. Dacheux, R. F., Frumkes, T. E. & Miller, R. F. *Brain Res.* **161**, 1-12 (1979).
12. Miller, R. F. & Dacheux, R. F. *J. gen. Physiol.* **67**, 639-659 (1976).
13. Werblin, F. S. in *The Neurosciences: Fourth Study Program* (eds Schmitt, F. O. & Worden, F. G.) 193-226 (MIT Press, 1979).
14. Olney, J. W. *J. Neuropath. exp. Neurol.* **28**, 455-474 (1969).
15. Schwarcz, R. & Coyle, J. T. *Invest. Ophthalmol.* **16**, 141-148 (1977).
16. Yazulla, S. & Kleinschmidt, J. *Brain Res.* **182**, 287-301 (1980).
17. Ikeda, H. & Sheardown, M. *J. Neuroscience* **7**, 25-36 (1982).
18. Miller, R. F., Slaughter, M. M. & Dick, E. in *Neurotransmitter Interaction and Compartmentation* (ed. Bradford, H. F.) 735-759 (Plenum, New York, 1982).
19. Slaughter, M. M. & Miller, R. F. *Science* **211**, 182-185 (1981).
20. Wunk, D. F. & Werblin, F. S. *J. gen. Physiol.* **73**, 265-286 (1979).
21. Davies, J., Evans, R. H., Francis, A. A., Jones, A. W. & Watkins, J. C. *J. Neurochem.* **36**, 1305-1307 (1981).
22. Slaughter, M. M. & Miller, R. F. *Science* **219**, 1230-1232 (1983).
23. Dowling, J. E. *Invest. Ophthalmol.* **9**, 655-680 (1970).
24. Kolb, H. *J. Neurocytol.* **8**, 295-329 (1979).
25. Miller, R. F., Frumkes, T. E., Slaughter, M. M. & Dacheux, R. F. *J. Neurophysiol.* **5**, 764-782 (1981).
26. Ikeda, H. & Sheardown, M. *J. Neuroscience* **7**, 1299-1308 (1982).
27. Masland, R. H. & Mills, J. W. *J. Cell Biol.* **83**, 159-178 (1979).
28. Slaughter, M. M. & Miller, R. F. *J. Neurosci.* (in the press).
29. Ishida, A. T. & Fain, G. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5890-5894 (1981).
30. Lasater, E. M. & Dowling, J. E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 936-940 (1982).

100 Hz remains upper limit of synchronous muscle contraction—an anomaly resolved

David Spencer Smith

Department of Zoology and Hope Entomological Collections,
University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Wootton and Newman¹ introduced a problem when they showed, by high-speed cinematography, that a minute hemipteran insect, the whitefly *Trialeurodes vaporariorum*, achieves a wing-beat frequency of up to 181 Hz. 100 Hz had been regarded²⁻⁴ as the upper limit of contraction by myoneural synchrony, and whitefly flight muscle had been placed in the 'synchronous' category on structural grounds⁵. The problem is resolved by evidence presented here, which reclassifies the power-producing flight muscles of whiteflies as 'asynchronous'.

While fast-acting vertebrate skeletal muscles and most insect muscles show synchrony between motor excitation and length change, it is well known^{2,3} that power-producing flight muscles, in some insect orders, have evolved the ability to dissociate these events allowing, though not necessitating, oscillatory con-

traction at frequencies—up to 1,000 Hz (ref. 6)—far beyond synchronous control of myofibrillar activation via calcium shuttling between the sarcoplasmic reticulum and the contractile system³. Synchronous and asynchronous muscles are readily distinguished by their fine structure^{7,8}: notably, the former possess an extensive sarcoplasmic reticulum compartment, associated with the T-system typically in triads (vertebrates) or dyads (arthropods). The sarcoplasmic reticulum volume varies with the rapidity of the activity cycle, and is hypertrophied in muscles operating synchronously at ~100 Hz (ref. 9). Plasma membrane invaginations of the T-system are present in asynchronous (fibrillar) insect muscles, but are associated in dyads with a very reduced sarcoplasmic reticulum^{3,7,8}; functionally interpretable in cells in which contraction is initiated as in synchronous muscle but in which oscillatory contraction is 'myogenic'³.

The structure of principal and ancillary thoracic muscles in three whiteflies (*Dialeurodes citrifolii*, *Aleurodes brassicae* and *T. vaporariorum*) was examined by conventional electron microscopy, the last being the species on which the present problem was based (Fig. 1a-c). In each case, the dorsal longitudinal flight muscle shows the structural features of other asynchronous fibres: the myofibrils are large, discrete and cylindrical (Fig. 1a), and the T-tubules accompany the fibrils at the mid-sarcomere level and there form dyads with the sarcoplasmic reticulum, very sparsely represented elsewhere (Fig. 1a, b). In contrast, small ancillary thoracic muscles (Fig. 1c) show exten-

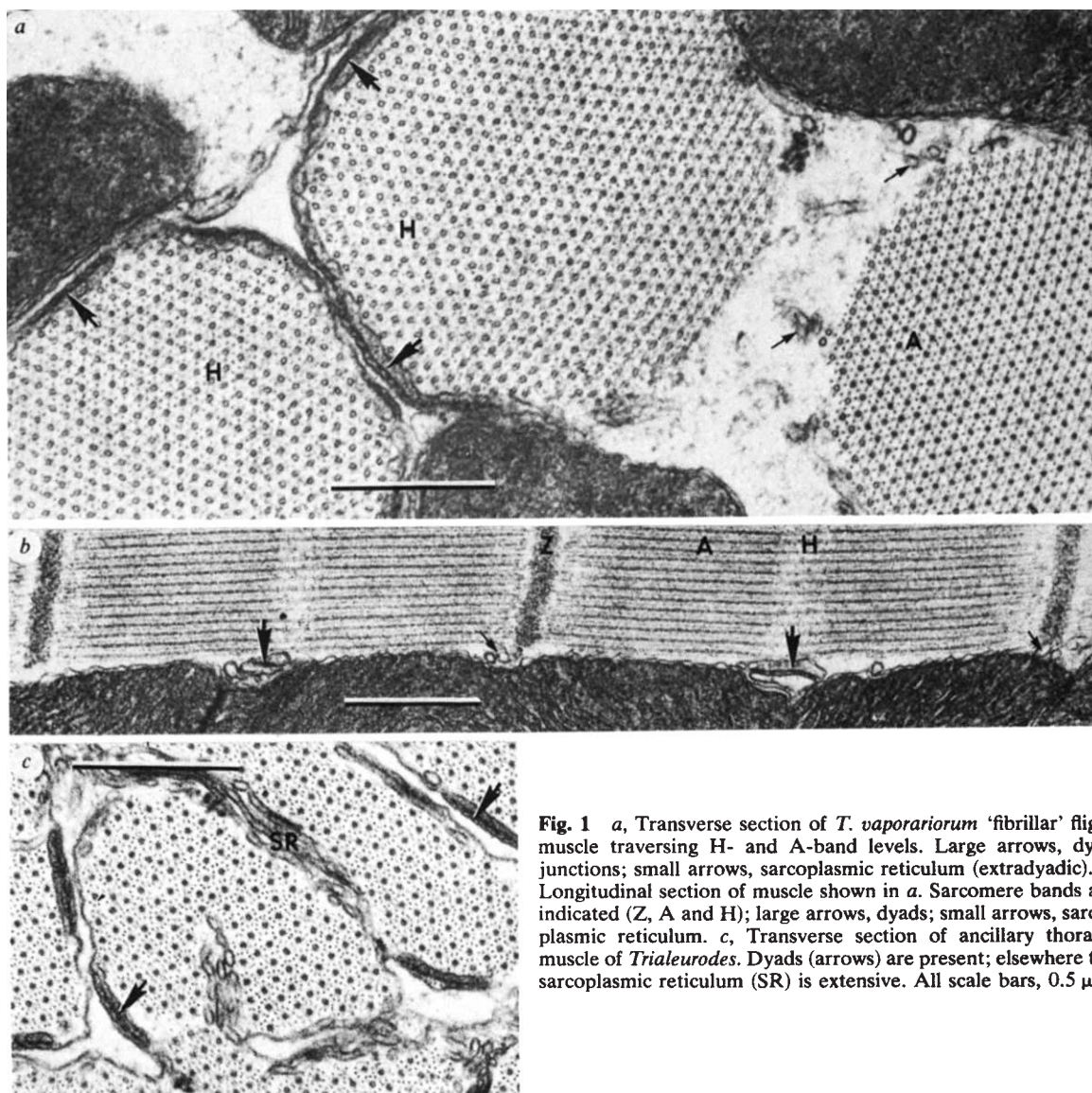


Fig. 1 a, Transverse section of *T. vaporariorum* 'fibrillar' flight muscle traversing H- and A-band levels. Large arrows, dyad junctions; small arrows, sarcoplasmic reticulum (extradyadic). b, Longitudinal section of muscle shown in a. Sarcomere bands are indicated (Z, A and H); large arrows, dyads; small arrows, sarcoplasmic reticulum. c, Transverse section of ancillary thoracic muscle of *Trialeurodes*. Dyads (arrows) are present; elsewhere the sarcoplasmic reticulum (SR) is extensive. All scale bars, 0.5 μ m.

sive sarcoplasmic reticulum and are clearly synchronous: as in other insects with an asynchronous flight mechanism⁵ these small fibres may adjust posture during flight and/or represent leg muscles, incorporated into the flight system.

Recognition of asynchrony in whitefly flight muscle is consistent with data on other Hemiptera, and on flight muscle of other very small insects, in which aerodynamic problems are met by high wingbeat frequency and special wing postures during the cycle¹⁰. Whiteflies include the smallest Hemiptera (wing and thoracic lengths in *Trialeurodes* are ~1 mm and 0.25 mm, respectively) and the recorded 181 Hz exceeds the frequency of other Homoptera⁵ and also of the Heteroptera which have uniformly evolved flight muscle asynchrony⁵. Wootton and Newman note that the observed frequency of *Trialeurodes* is unexpectedly low for an insect of this size, and suggest that both the 'clap and fling' wing posture¹⁰ and unusually low wing loading contribute to flight efficiency in these insects.

In the context of evolution of asynchrony in the Homoptera⁵, relocation of the whiteflies leaves six rather than seven of 10 families in the 'synchronous' list. To muscle physiologists, the correction is of more general significance, in reinstating ~100 Hz as the probable limit of myoneural synchrony.

Received 21 February; accepted 13 April 1983.

1. Wootton, R. J. & Newman, D. J. S. *Nature* **280**, 402-403 (1979).
2. Pringle, J. W. S. *Symp. R. ent. Soc. Lond.* **7**, 3-15 (1976).
3. Pringle, J. W. S. *J. exp. Biol.* **94**, 1-14 (1981).
4. Mendelson, M. J. *Cell Biol.* **42**, 548-561 (1969).
5. Cullen, M. J. *J. Ent.* **A49**, 17-41 (1974).
6. Sotavalta, O. *Acta ent. fenn.* **4**, 1-117 (1947).
7. Smith, D. S. *J. biophys. biochem. Cytol.* **10**, Suppl. 123-158 (1961).
8. Smith, D. S. *Prog. Biophys. molec. Biol.* **16**, 107-142 (1966).
9. Rosenbluth, J. J. *Cell Biol.* **42**, 534-547 (1969).
10. Weis-Fogh, T. in *Scale Effects in Animal Locomotion*, 405-420 (Academic, New York, 1977).

Z-DNA immunoreactivity in rat tissues

Göpf Morgenegg*, Marco R. Celio†, Bernard Malfoy‡, Marc Leng‡ & Clive C. Kuenzle*

* Institut für Pharmakologie und Biochemie, Veterinärmedizinische Fakultät der Universität Zürich, Winterthurerstrasse 260, CH-8057 Zürich, Switzerland

† Anatomisches Institut der Universität Zürich, Gloriastrasse 19, CH-8006 Zürich, Switzerland

‡ Centre de Biophysique Moléculaire, CNRS, 1A, avenue de la Recherche Scientifique, F-45045 Orléans Cedex, France

Recently, it has been shown that natural and synthetic deoxynucleotide polymers can adopt a left-handed helical structure (termed Z-DNA) in appropriate conditions (see, for example, refs 1 and 2 and the references therein). In contrast to the more familiar right-handed B-DNA, Z-DNA is strongly immunogenic, and polyclonal and monoclonal antibodies against Z-DNA have been elicited³⁻⁶. By using such antibodies, immunoreactivity for Z-DNA has been detected in the polytene chromosomes of two dipteran species⁷⁻⁹, in the macronucleus of a ciliated protozoan¹⁰, and in certain plant nuclei (cited in ref. 11). In view of the possible importance of Z-DNA as a genomic regulatory signal⁷, it would be highly desirable to know whether Z-DNA also occurs in mammals. We have therefore initiated an immunohistochemical study of various rat tissues by using three antisera specific for Z-DNA, and the peroxidase-antiperoxidase technique¹² for visualization of tissue-bound antibodies. Here we demonstrate that the nuclei of many, but not all, types of rat cells exhibit Z-DNA immunoreactivity, suggesting that Z-DNA may exist naturally in mammalian chromatin.

The tissue distribution of Z-DNA immunoreactivity was investigated using three rabbit antisera directed against Z-

DNA. Two of these (α -Z1 and B13) were elicited by immunizing animals with poly(dG-dC)·poly(dG-dC) modified on 12% of the bases with chloro(diethylenetriamine)platinum(II) chloride⁴ [this modification will hereafter be abbreviated as dienPt(0.12)]. The third antiserum (α -Z2) was raised against poly(dG-m⁵dC)·poly(dG-m⁵dC)dienPt(0.12). Both the unmethylated and the methylated Pt-modified polynucleotides exist in the Z-form in physiological conditions (refs 4, 13 and our own unpublished results). The detailed characterization of α -Z1 has been published^{4,14,15}, and qualitatively similar results were obtained with the other two antisera (unpublished).

All three antisera are specific for Z-DNA, reacting with several Z-form polynucleotides, such as the high-salt form of poly(dG-dC)·poly(dG-dC), poly(dG-dC)·poly(dG-dC)dienPt(0.12), poly(dG-m⁵dC)·poly(dG-m⁵dC)dienPt(0.12) and poly(dG-Br⁵dC)·poly(dG-Br⁵dC) (a brominated derivative stably fixed in the Z-conformation¹⁴). In contrast, the antisera do not recognize double-stranded synthetic polynucleotides or natural and modified DNAs in either the Z or B-forms, relaxed double-stranded plasmid DNA, single-stranded natural and modified DNA, double-stranded RNA, poly(dG)·poly(dC), poly(dG), deoxyguanosine or deoxycytidine. The three antisera can therefore be used to search for Z-DNA by immunohistochemistry. Experiments were also conducted with affinity-purified antibodies isolated from one of these antisera (α -Z1) by chromatography on Sepharose-poly(dG-dC)·poly(dG-dC) in high salt conditions¹⁴. The tissue-bound antibodies were traced by means of the unlabelled antibody technique involving the peroxidase-antiperoxidase complex¹².

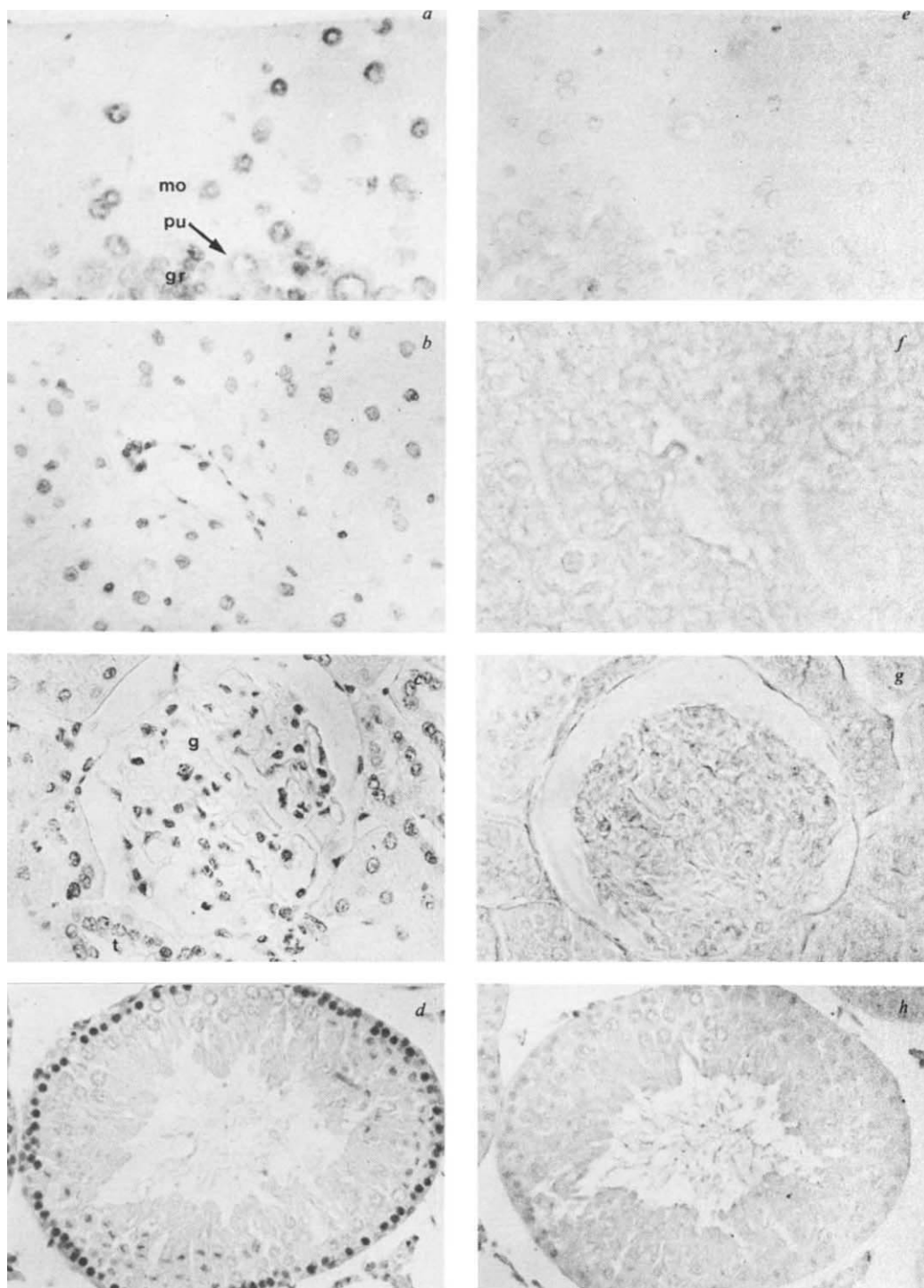
Figure 1a-d shows the distribution of Z-DNA immunoreactivity in four tissues of the rat as detected with the antiserum α -Z1. Qualitatively the same results were obtained with antisera B13 and α -Z2, as well as with the affinity-purified antibodies (not shown). In contrast, an almost complete absence of staining was observed with a non-immune serum used as a control (Fig. 1e-h). Control reactions were equally negative with two other non-immune sera and normal rabbit immunoglobulin G, the latter being applied at a concentration matching that of the affinity-purified antibodies (not shown). The panels represent cerebellar cortex (Fig. 1a, e), liver (Fig. 1b, f), kidney (Fig. 1c, g) and testis (Fig. 1d, h). It can be seen that most, but not all, nuclei exhibit Z-DNA immunoreactivity. This is most clearly exemplified in the testis (Fig. 1d, h), where the darkly stained nuclei of spermatogonia are aligned to form a prominent circle at the periphery of the seminiferous tubules, whereas the nuclei of spermatocytes, spermatids and sperm, located closer to the lumen, remain largely unstained. Another example is found in the cerebellar cortex (Fig. 1a, e), where strongly stained and virtually unstained nuclei coexist in the molecular layer (mo). The cerebellar cortex also demonstrates that immunoreactive nuclei are not all stained to the same degree. Compared with the most intensely stained nuclei in the molecular layer (mo) those in the granule cell layer (gr) are more weakly stained, and the immunoreactivity of Purkinje neurones (pu) is questionable. The kidney sections (Fig. 1c, g) also give the impression of more strongly reactive nuclei in the glomerula (g) than in the tubuli (t).

In all tissues examined only the nuclei were stained, which is consistent with the location of the bulk of the DNA in mammalian cells in the nuclei. Except in the spermatogonia, where the staining was uniform (Fig. 1d), the immunoreaction appeared to be restricted to distinct regions of the nucleus (Fig. 1a-c). This may indicate that chromatin consists of alternating regions enriched for either Z-DNA or B-DNA.

To confirm that the immunoreactivity in the nuclei really was due to the presence of Z-DNA, the antisera were absorbed with a 40-fold excess of various polynucleotides. The ratio of polynucleotide to antiserum was calculated from previous determinations, which had shown that the antisera contain 0.45(α -Z1), 0.27(B13) and 0.1(α -Z2) mg of Z-DNA-specific IgG per ml of serum, and that each antibody binding site covers approximately four nucleotide residues (ref. 14 and unpublished

Fig. 1 Immunohistochemistry for Z-DNA in four tissues of the rat: cerebellar cortex (*a, e*), liver (*b, f*), kidney (*c, g*) and testis (*d, h*). Fixation was with 45% acetic acid. Sections *a-d* were treated with antiserum α -Z1 directed against Z-DNA, and sections (*e-h*) with a non-immune serum for control. The sera were applied at a dilution of 1:5,000. The peroxidase-antiperoxidase (PAP) technique was used for visualization of the primary antibodies bound to the tissue. Primary magnification $\times 400$, except for the testis, which is $\times 200$. Abbreviations: mo, molecular layer; gr, granule cell layer; pu, Purkinje neurones; g, glomerula; t, tubuli.

Methods: Male Zur:SVZ rats, 37-40 weeks old, were killed by decapitation and the organs removed immediately. Fixation was with 45% (v/v) acetic acid for 8 h at room temperature. The tissues were embedded in paraffin, and 5 μ m-thick sections were collected on chromalum gelatine-coated glass slides. The sections were dewaxed with xylene and rehydrated. They were incubated for 72 h at 4 °C in a moist chamber with the primary antiserum diluted 1:1,000 to 1:10,000 in PBS (phosphate-buffered saline consisting of 137 mM NaCl, 3 mM KCl, 10 mM Na/K-phosphate, pH 7.4). The primary antibody bound to the tissue was traced by means of the unlabelled antibody technique involving the PAP complex¹². Goat anti-rabbit IgG (Nordic, Copenhagen; 1:50 in PBS) and the PAP complex (Sternberger-Meyer Inc., Jarrettsville Pike, Maryland; 1:200 in buffer composed of 150 mM NaCl, 1 mM EDTA, 0.3 mM NaN₃, 50 mM Tris-HCl pH 7.4, 0.1% (w/v) bovine serum albumin and 0.05% (v/v) Tween 20) were each applied for 30 min at room temperature. The location of the antibody-bound peroxidase was then visualized by incubation with the substrate 3,3'-diaminobenzidine-4 HCl/hydrogen peroxide. Each of the various steps was followed by two washes (5 min) with PBS. The tissue sections were then dehydrated and mounted.



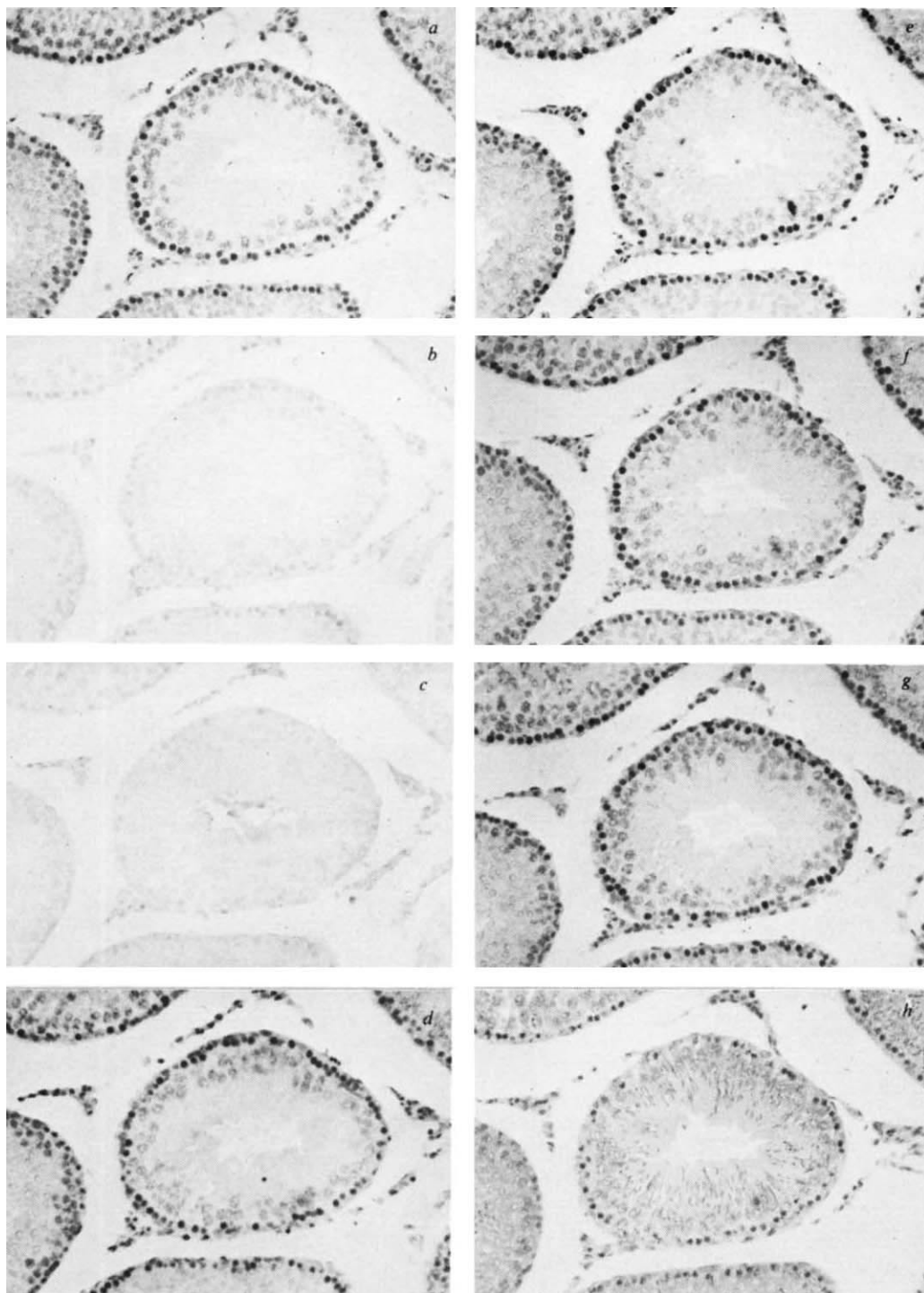
results). Accordingly, absorptions were performed at ratios of 333, 200 and 75 ng polynucleotide per μ l of serum, respectively.

Figure 2 summarizes the results of the absorption experiments performed on consecutive sections of testis tissue. The reactions presented are those elicited with the antiserum α -Z1, but the patterns are representative not only of this antiserum but also of B13 and α -Z2 (not shown). Results analogous to those with testis were also obtained for cerebellum, liver and kidney (not shown). Figure 2a represents a sham absorption with buffer only. Again, the nuclei of the spermatogonia stand out as a prominent ring giving a picture indistinguishable from that seen with the untreated serum sample (Fig. 1d). In contrast,

immunostaining was clearly abolished following absorption of the antiserum with poly(dG-dC).poly(dG-dC)dientPt(0.12) (Fig. 2b) and with poly(dG-Br⁵dC).poly(dG-Br⁵dC) (Fig. 2c), both of which adopt the Z-conformation in the low-salt conditions of absorption^{4,13,14}. No reduction in immunostaining ensued from treating the antiserum with B-form poly(dG-dC).poly(dG-dC) (Fig. 2d), calf thymus double-stranded DNA (Fig. 2e), calf thymus single-stranded DNA (Fig. 2f) or calf liver RNA (Fig. 2g). A conventionally stained tissue section is shown for comparison in Fig. 2h. The combined results demonstrate that the antisera specifically recognized Z-DNA in the tissues investigated.

Fig. 2 Immunohistochemistry for Z-DNA in the rat testis after absorption of antiserum α -Z1 with various polynucleotides in physiological conditions. Absorbents were: pure buffer without polynucleotide (a); poly(dG-dC) · poly(dG-dC)-dienPt(0.12) (b); poly(dG-Br⁵dC) · poly(dG-Br⁵dC) (c); poly(dG-dC) · poly(dG-dC) (d); double-stranded calf thymus DNA (e); single-stranded calf thymus DNA (f); and calf liver RNA (g). A conventional stain with haematoxylin-eosin is shown for comparison (h). All absorptions were done at a ratio of 333 ng polynucleotide per μ l of antiserum, which constitutes a 40-fold antigen excess. The serum was diluted 1:5,000 and was applied to consecutive tissue sections fixed with 45% acetic acid. The PAP technique was used for visualization of the primary antibodies bound to the tissue. $\times 160$.

Methods: All operations were carried out at 4 °C. For absorption, 1 μ g polynucleotide in 5 ml buffer consisting of 150 mM NaCl, 1 mM MgCl₂, 50 mM Tris-HCl, pH 7.5, was gently mixed with 3 μ l antiserum and incubated for 24 h in a polyallomer tube. The mixture was centrifuged in a Beckman SW 50.1 rotor at 105,000 g for 20 min, and the supernatant was stored at -80 °C until use. It was then diluted with the same buffer for incubation with the tissue sections. Otherwise, immunohistochemistry was as described in Fig. 1 legend. Poly(dG-dC) · poly(dG-dC)-dienPt(0.12) and poly(dG-Br⁵dC) · poly(dG-Br⁵dC) were prepared as described in refs 13 and 14, respectively. Calf liver RNA was from Sigma and all other polynucleotides were from PL-Biochemicals.



Tissue fixation is a well recognized source of artefacts in immunohistochemistry. We relied on fixed tissue sections in the present investigation, and to minimize possible objections against this approach several fixatives were explored for their suitability. Apart from acetic acid (Figs 1 and 2) three other fixatives were found to give positive results (Fig. 3): these were picric acid-HgCl₂ (Fig. 3a), methanol (Fig. 3b) and acetone applied to the tissue by freeze-substitution (Fig. 3c). The same qualitative pattern of Z-DNA immunoreactivity was detected in the testis irrespective of the method of fixation. Only the intensity of the immunochemical staining and the degree of tissue preservation were influenced by the choice of the fixative. In both respects, acetic acid gave the best results, so most experiments were done using this fixative. Whether acetic acid

is able to convert natural or synthetic DNA to the Z-form is unknown; however, one could imagine a mechanism by which acetic acid might induce the formation of Z-DNA. Kuo¹⁶ has shown that acetic acid destroys the nucleosome structure in chromatin. This may bring about the uncoiling of the two turns of DNA wrapped around each nucleosome core, generating a vast amount of superhelicity¹⁷ and sufficient torsional strain to force the DNA into the Z-conformation^{5,15,18,19}. A similar mechanism could also be invoked for picric acid, but it is less likely that fixation with methanol and acetone would involve nucleosome destruction. Whether such a mechanism actually operates is difficult to test at the present, but it seems unlikely that four unrelated fixatives could induce the same type of artefact with equal distribution in nuclei. This supports the

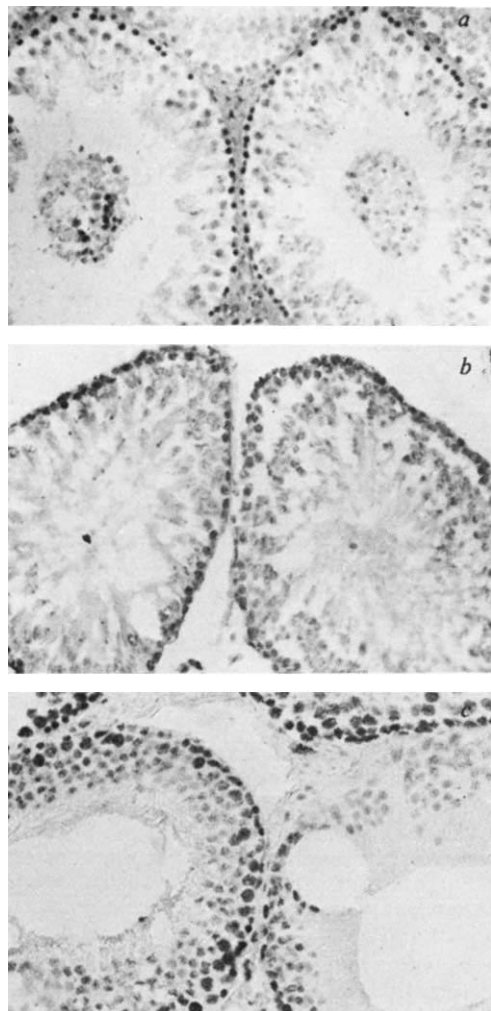


Fig. 3 Effect of tissue fixation on Z-DNA immunoreactivity in the rat testis. Fixation was with picric acid-HgCl₂ (a), 50% methanol (b) or by freeze-substitution with acetone (c). Antiserum α -Z1 was applied at a dilution of 1:5,000 (a, b) or 1:10,000 (c). The PAP technique was used for visualization of the primary antibodies bound to the tissue. $\times 160$.

Methods: Fixation with picric acid-HgCl₂ (saturated aqueous picric acid-saturated aqueous HgCl₂-water, 1:1:2 by volume) was done by immersion of tissue for 8 h at room temperature. Fixation with 50% (v/v) methanol (250 ml) involved perfusion of the animal through the aorta. Freeze-substitution with acetone was performed for 1 week at -70°C and was followed by further fixation at the same temperature with a mixture of acetone (90 ml), 40% formalin (5 ml), acetic acid (1 ml) and water (4 ml) for another week; after thawing the specimens were embedded in paraffin.

All other conditions as in Fig. 1.

notion that Z-DNA is a natural element in mammalian chromatin.

The Z-conformation was originally observed with deoxynucleotide polymers of strictly alternating purine-pyrimidine sequences. With polymers of the type poly(dG-dC).poly(dG-dC) and related sequences, the Z-form occurs both in the solid state and in solution^{1,2}. It is more difficult to convert poly(dA-dC).poly(dG-dT) from the B- to the Z-form²⁰⁻²³, and it seems that poly(dA-dT).poly(dA-dT) probably cannot acquire the left-handed conformation. The B $\rightarrow$ Z transition can be brought about by changes in solvent composition, by covalent modification with bulky substituents and by torsional strain induced by supercoiling^{1,2}. The last of these mechanisms may be relevant to the formation of Z-DNA *in vivo*. Supercoiling is a well recognized phenomenon in living systems and long stretches of poly(dA-dC).poly(dG-dT)²⁴⁻²⁶ and of poly(dG-dC).poly(dG-dC)²⁶ have been detected as repetitive elements

in mammalian DNA. However, long alternating sequences may not be necessary for Z-DNA formation in natural DNA in view of the finding that circular, covalently closed plasmid DNA, devoid of extended alternating sequences, can be induced to adopt the Z-conformation when put under torsional stress^{5,15,18,19}. Alternatively, the generation of Z-DNA *in vivo* may be mediated by chromosomal proteins binding to Z-DNA specifically. In fact, Z-DNA-binding proteins have been detected in *Drosophila* nuclei¹¹, and recent unpublished evidence from our laboratories suggests that similar proteins also exist in the cell nuclei of rat and bovine brain.

We thank Mrs Helga Weber for microphotography and Miss Christina Hemmerle for technical assistance. This work was supported by Swiss NSF grant 3.004-0.81 and by the French Institut National de la Santé et de la Recherche Médicale (contrat no. 120019).

Received 3 February; accepted 7 April 1983.

1. Zimmerman, S. B. A. *Rev. Biochem.* **51**, 395-427 (1982).
2. Neidle, S. *Nature* **302**, 574 (1983).
3. Lafer, E. M., Möller, A., Nordheim, A., Stollar, B. D. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3546-3550 (1981).
4. Malfoy, B. & Leng, M. *FEBS Lett.* **132**, 45-48 (1981).
5. Pohl, F. M., Thomae, R. & DiCapua, E. *Nature* **300**, 545-546 (1982).
6. Möller, A. *et al. J. biol. Chem.* **257**, 12081-12085 (1982).
7. Nordheim, A. *et al. Nature* **294**, 417-422 (1981).
8. Lemeunier, F., Derbin, C., Malfoy, B., Leng, M. & Taillandier, E. *Expl Cell Res.* **141**, 508-513 (1982).
9. Leng, M. *et al. Cold Spring Harb. Symp. quant. Biol.* **47** (in the press).
10. Lipps, H. J. *et al. Cell* **32**, 435-441 (1983).
11. Nordheim, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7729-7733 (1982).
12. Sternberger, L. A. *Immunocytochemistry* 2nd edn (Wiley, New York, 1979).
13. Malfoy, B., Hartmann, B. & Leng, M. *Nucleic Acids Res.* **9**, 5659-5669 (1981).
14. Malfoy, B., Rousseau, N. & Leng, M. *Biochemistry* **21**, 5463-5467 (1982).
15. Lang, M. C., Malfoy, B., Freund, A. M., Daune, M. & Leng, M. *EMBO J.* **1**, 1149-1153 (1982).
16. Kuo, M. T. *Expl Cell Res.* **138**, 221-229 (1982).
17. Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843-1847 (1975).
18. Brahms, S. *et al. J. molec. Biol.* **162**, 473-493 (1982).
19. Nordheim, A. *et al. Cell* **31**, 309-318 (1982).
20. Arnott, S., Chandrasekaran, R., Birdsall, J. J., Leslie, A. G. W. & Ratliff, R. L. *Nature* **283**, 743-745 (1980).
21. Wells, R. D. *et al. J. biol. Chem.* **257**, 10166-10171 (1982).
22. Vorlíčková, M., Kypr, J., Štokrová, Š. & Šponar, J. *Nucleic Acids Res.* **10**, 1071-1080 (1982).
23. Zimmer, C., Tymen, S., Marck, C. & Guschi, W. *Nucleic Acids Res.* **10**, 1081-1091 (1982).
24. Nishioka, Y. & Leder, P. *J. biol. Chem.* **255**, 3691-3694 (1980).
25. Hamada, H. & Kakunaga, T. *Nature* **298**, 396-398 (1982).
26. Hamada, H., Petrino, M. G. & Kakunaga, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6465-6469 (1982).

Pb(II)-catalysed cleavage of the sugar-phosphate backbone of yeast tRNA^{Phe}—implications for lead toxicity and self-splicing RNA

R. S. Brown*, B. E. Hingerty*, J. C. Dewan & A. Klug

MRC Laboratory of Molecular Biology, MRC Centre, University Medical School, Hills Road, Cambridge CB2 2QH, UK

Pb(II) is extremely efficient at depolymerizing RNA¹⁻⁵ and studies on tRNAs have shown that site-specific cleavages in these molecules can be brought about by the action of Pb(II)⁶. We have observed, by difference Fourier analysis^{7,8}, sugar-phosphate strand scission between residues 17 and 18 in crystals of yeast tRNA^{Phe} soaked in dilute Pb(II) solution at pH 7.4. We have also deduced the structure of the Pb(II)-tRNA^{Phe} derivative at pH 5.0 where this cleavage reaction is considerably slower and report that, in this structure, the sugar-phosphate backbone remains intact. We have, therefore, a picture of the reactants (at pH 5.0) and products (at pH 7.4) of this cleavage reaction. From this crystallographic study, and associated biochemical work, we have formulated a possible mechanism for the cleavage reaction and also present here some general ideas on the action of metal ions on nucleic acids.

* Present addresses: European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, FRG (R.S.B.); Oak Ridge National Laboratory, Health and Safety Research Division, PO Box X, Oak Ridge, Tennessee 37830, USA (B.E.H.).

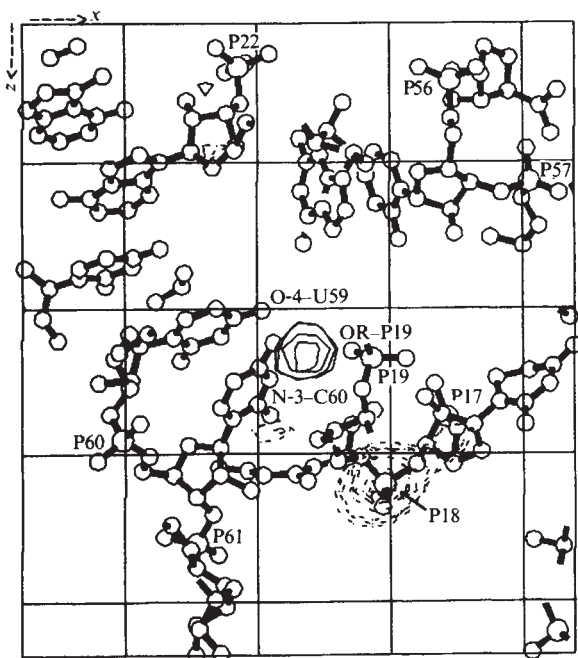


Fig. 1 View of the $|F_{\text{Pb}}| - |F_{\text{Native}}|$ difference Fourier map of yeast tRNA^{Phe} at pH 7.4 in a slab 8 Å thick centred on the Pb(II) ion at the major cleavage site Pb(1). The refined native structure is superimposed and the phosphorus atoms are labelled. Contours are ± 0.4 , ± 0.8 and $\pm 1.2 \text{ e}\text{\AA}^{-3}$ with negative contours shown as broken lines. The large positive peak corresponds to the Pb(II) ion at site Pb(1) while the large negative peak, centred on phosphate 18, represents sugar-phosphate backbone cleavage between residues D17 and G18. The Pb(II) ion is bound directly to atom O-4 of the base U59 and to N-3 of C60 with a long contact to oxygen OR of phosphate 19.

X-ray data were collected at -5°C from monoclinic yeast tRNA^{Phe} crystals soaked at 4°C in 1 mM Pb(II) acetate where the Pb(II)/tRNA^{Phe} ratio was 5:1. The experiment was carried out at pH 5.0 and pH 7.4 to a resolution of 3 Å in each case. Both Pb(II) derivative crystals are isomorphous with the native monoclinic crystals whose structure, refined at 2.5 Å resolution, has been reported previously^{9,10}. For both of the derivative structures, difference Fourier maps computed using coefficients $|F_{\text{Pb}}| - |F_{\text{Native}}|$ and native phases revealed three Pb(II) binding sites which have essentially the same coordinates at either pH. These coordinates and the occupancies for the Pb(II) sites are given in Table 1. The difference map at pH 5.0 was featureless apart from three positive peaks corresponding to the three Pb(II) ions, indicating that there is good isomorphism between the native and derivative structures, and that the tRNA structure has not been changed markedly by the Pb(II) binding. In the difference map at pH 7.4 we observed the same three positive peaks due to Pb(II) ions, and a significant negative peak centred on the position of phosphate 18 (Fig. 1). These results were interpreted to mean that at pH 7.4, significant sugar-phosphate backbone cleavage of the tRNA^{Phe} molecule had occurred between residues D17 and G18, whereas at pH 5.0 the cleavage reaction had not proceeded to any observable extent (note that D=dihydrouracil, G=guanine and the numerals represent the numbering in the sequence of the tRNA⁹).

Cleavage of the tRNA^{Phe} molecule, both in solution and in crystals, was also followed by using polyacrylamide gel electrophoresis. After soaking tRNA^{Phe} crystals at pH 7.4 in 1 mM Pb(II) solution at 4°C for 1 week, two cleavages were evident, one of which was a major and the other a minor cleavage site. Further biochemical investigation has shown that the major cleavage corresponds to that between residues D17 and G18, the cleavage observed in the difference maps at this pH. The minor cleavage was between G22 and A23. At pH 5.0,

Table 1 Crystallographic coordinates (in Å) and occupancies (in electrons) of the three Pb(II) ions in the two crystal structures of the Pb(II)-tRNA^{Phe} derivatives

	x	y	z	Occupancy*
Pb(1)	52.85 (52.62)	10.80 10.91	32.70 32.62	24 (50)
Pb(2)	46.46 (46.45)	8.57 8.43	16.40 16.26	34 (40)
Pb(3)	44.45 (44.50)	-10.28 -10.36	-1.21 -1.20	61 (63)

Values for the structure at pH 5.0 appear (in parentheses) below those for the structure at pH 7.4.

* A value of 80 represents an occupancy of 100%.

again after soaking for 1 week in a 1 mM Pb(II) solution at 4°C , cleavage at the major site between residues D17 and G18 was substantially reduced. Historically, work on the structure at pH 7.4 was performed first as part of a search for heavy-atom derivatives. After observing strand cleavage in the crystals at this pH, we surmised, as later justified by the electrophoretic analysis outlined above, that at pH 5.0 we should be able to collect X-ray data before significant strand cleavage had occurred, thus giving a picture of the Pb(II) binding before cleavage. The four crystals used for the X-ray experiment at pH 5.0 were each soaked in Pb(II) solution for only 24 h before the start of the X-ray data collection and were discarded after 6 days on the diffractometer, that is, before any significant reaction could occur. We believed that the location of the Pb(II) ions before and after strand scission would be important in formulating a mechanism for the observed cleavage.

The three Pb(II) binding sites have been designated Pb(1), Pb(2) and Pb(3). At site Pb(1) a Pb(II) ion is bound, in the single-stranded T ψ C loop region of the tRNA^{Phe} molecule, to atom O-4 of the base U59 at a distance of 2.2 Å, to N-3 of C60 at 2.8 Å, with a long contact to oxygen OR of phosphate 19 of 3.4 Å (Fig. 1). The Pb(II) ion bound at site Pb(1) displaces the Mg(II) ion bound near this site in the native structure, designated Mg(3) in ref. 7. Figure 11 of ref. 7 gives an overall view of the various metal-binding sites in the tRNA^{Phe} structure. At site Pb(2) a second Pb(II) ion is bound in the extra loop region to O-6 of G45 at a distance of 2.5 Å with the Pb(2) ... N-7 (G45) distance being 3.6 Å. The Pb(II) ion bound at this site displaces a spermine molecule and the site is the same as that labelled Sm(4) in ref. 7. The third Pb(II) ion, Pb(3), is bound directly to N-7 of Y37 in the anticodon region at a distance of 2.4 Å and displaces what has been assigned as a water molecule in the native monoclinic structure. Site Pb(1) is close to the major cleavage between residues D17 and G18, while the minor cleavage site, between G22 and A23, is near site Pb(2). No sugar-phosphate backbone cleavage is observed near site Pb(3).

The major strand cleavage occurs between residues D17 and G18 which, because of the three-dimensional conformation of the tRNA^{Phe} molecule, are near the Pb(1) site. We propose that this cleavage occurs via a Pb(II)-bound hydroxyl group¹¹ at site Pb(1) that abstracts the proton from the 2'-OH of ribose 17, thus facilitating nucleophilic attack of the resulting 2'-O⁻ on the phosphorus atom of phosphate 18 (Fig. 2)¹². The reaction will proceed through a pentacoordinate phosphorus intermediate to produce a 2',3'-cyclic phosphate at ribose 17 (ref. 12). The end group at residue G18 will be a 5'-OH. We have demonstrated that we do indeed have these end groups at this cleavage site by extensive biochemical analysis. This analysis also indicates that >90% of the cyclic phosphate remains intact. The distance between the Pb(II) ion at site Pb(1) and the oxygen atom of the 2'-OH of ribose 17 is 6.0 Å, which is appropriate for the proposed interaction between Pb(II) and ribose 17 via an intervening hydroxyl group. Pb(II)-O(water) distances¹³ can be as large as ~ 3 Å and a similar distance for the O ... O interaction also seems reasonable. At a resolution of 3 Å we are unable to locate water molecules bound to the Pb(II) ions.

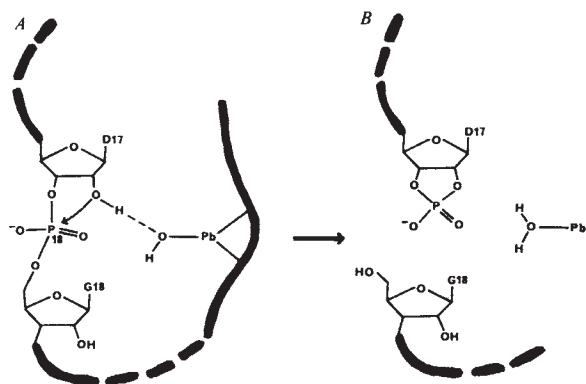


Fig. 2 Diagrammatic representation of the proposed mechanism for the sugar-phosphate backbone cleavage between residues D17 and G18 in monoclinic crystals of yeast tRNA^{Phe}. In *A* (before cleavage of the sugar-phosphate backbone) the Pb(II) moiety is depicted as being bound to, and positioned by, the tRNA^{Phe} molecule, while in *B* (after strand cleavage) the Pb(II) moiety is depicted as being free of the tRNA^{Phe} molecule.

The fact that there are only two direct bonding contacts that we can detect between the Pb(II) ion at site Pb(1) and the tRNA^{Phe} molecule strongly suggests, however, that the metal coordination sphere must, to some extent, be filled by such solvent-derived ligands. The mechanism proposed above is similar to that observed for the first step of the reaction in ribonuclease¹⁴⁻¹⁶, although examination of the tRNA^{Phe} structure and molecular models suggests that the present reaction may proceed via an adjacent mechanism rather than the in-line mechanism observed for ribonuclease.

In previous studies¹⁷⁻¹⁹ it has been suggested that an electrophilic activation mechanism¹¹ is responsible for the metal depolymerization of RNA. This may be so in other systems. However, an examination of Pb(1)...O(phosphate) distances shows that the electrophilic activation mechanism, whereby Pb(II) binds directly to a phosphate oxygen, polarizing the P—O bond and thus initiating nucleophilic attack of the 2'-OH on the phosphorus atom, is not possible in the present case. The closest such distance is Pb(1)...OR(phosphate 19) of 3.4 Å, which should lead to cleavage between residues G18 and G19; this cleavage is not observed. All remaining Pb(1)...O(phosphate) distances are greater than 4.8 Å. Moreover, the fact that the cleavage takes place rapidly at pH 7.4, close to the value for the pK_a of a water molecule bound to Pb(II)²⁰, and only slowly at pH 5.0, suggests that a hydroxyl group bound to the Pb(II) ion is involved in the nucleophilic attack.

The minor cleavage site between residues G22 and A23, detected by gel electrophoresis although not observed in the crystal structures, is near the Pb(2) site. We have, as yet, been unable to formulate a plausible mechanism for this cleavage reaction that fits both the structural and biochemical data although the fact that cleavage occurs near a Pb(II) ion strongly suggests metal ion involvement. At site Pb(3), in the anticodon region, no strand cleavage is observed.

Three factors seem to have combined fortuitously to effect the catalysis of the major strand cleavage at site Pb(1). First, the tRNA^{Phe} molecule in both native and derivative structures has very large temperature factors for the residues D16 and D17 (refs 9, 10), indicating that this section of the molecule is more flexible than the rest. Second, we presume that Pb(II) binds in such a manner that it presents a hydroxyl group in, or close to, the correct orientation and at the correct distance to effect the proton abstraction that we have suggested. Third, the pK_a of a Pb(II)-bound water molecule²⁰ falls around pH 7.4 (that is, physiological pH), thus facilitating reaction at this pH and slowing it substantially at pH 5.0. Most metal-bound water molecules possess a pK_a rather higher than physiological pH, for example, the value for Zn(II) is ~9.0, that for Mg(II) is ~12.0 while the lanthanides(III) have values of ~8.5 (ref. 20).

These metals will cleave RNA but require pH conditions far removed from physiological pH. In this respect Pb(II) is unusual although the values for Cu(II) and Ni(II) can also fall in this region. Note also that the coordination chemistry of Pb(II) is characterized by a wide variety of coordination numbers and geometries²¹. In this respect it is rather less rigorous in its coordination requirements than are many other metals and therefore may bind to folded RNA more easily.

The fact that the pK_a of a water molecule bound to Pb(II) falls in the region of physiological pH may have broad implications for lead toxicity via its known ability to rapidly depolymerize RNA. The reasons for lead toxicity are obscure and even its binding sites *in vivo* are not well documented compared with those of other toxic metals. Lead may exert its toxicity *in vivo* by its ability to rapidly cleave RNA, for example messenger RNA⁵, rather than by binding to sulphhydryl groups of proteins as is one common assumption.

We believe that the major Pb(II) cleavage observed in the tRNA^{Phe} crystals is catalytic and not stoichiometric, as the occupancy at site Pb(1) is reduced by ~50% in the structure at pH 7.4 compared with that observed in the structure at pH 5.0 (Table 1). This indicates that once the cleavage reaction has occurred, the Pb(II) ion is freed and so can move on to effect the same reaction in further tRNA^{Phe} molecules (Fig. 2). The Pb(II) occupancies at the other two sites are similar at either pH. Because the cleavage by Pb(II) is catalytic, there is no reason to expect a threshold effect. Indeed, work by ourselves and others^{5,6} shows that if sufficient time is allowed, cleavage will occur even at very low concentrations of Pb(II) and low Pb(II)/RNA ratios.

The work described here for yeast tRNA^{Phe} has wider implications for RNAs in general. The Pb(II) ions in the present case bind in complexly folded single-stranded regions, which is presumably one reason why single-stranded RNA is more easily degraded by metal ions than double-stranded RNA. Note also that the cleavage in the crystal takes place not in the most accessible single-stranded region of the molecule (that is, the CCA end), but in a region where the Pb(II) ion is presented with a variety of possible ligand arrangements. The major strand cleavage at the Pb(1) site occurs simply because several necessary conditions are fortuitously met simultaneously. At the minor cleavage site this is presumably not the case, resulting in slow cleavage, while at site Pb(3) no cleavage occurs, indicating that one or more of the necessary conditions is lacking. In this respect the Pb(II) ion at site Pb(1) acts very much like a local metallo-enzyme wherein rigorous stereochemical requirements are met so as to greatly accelerate a specific reaction. In the present case, Pb(II) bound to one strand of the tRNA^{Phe} cleaves another section of the molecule that passes nearby. It is the special configuration at site Pb(1) that allows the Pb(II) ion to cleave the tRNA so efficiently. While this particular site is adventitious in this particular molecule, there is every reason to believe that similar situations exist in other folded RNAs, such as in ribosomes. Indeed, almost any RNA molecule will tend to fold and provide pockets of the kind described, and it is probable that at least one of them will bind and position Pb(II) to act, as it were, as a local metallo-enzyme and so cause self-destruction.

Based on the observations presented here, it is also tempting to consider the possible relevance of metal ions in the recently reported phenomenon of self-splicing RNA, labelled the 'ribozyme'^{22,23}. Is it possible that site-bound metal ions, present in trace amounts or as added Mg(II) ions, for example, have an important chemical, as well as structural, role in this process?

More extensive details of the work reported here are being prepared for publication. A preliminary account of the crystallographic results at pH 7.4 was presented at the Boston Meeting of the American Crystallographic Association, August 1979. We thank Dr D. Rhodes for help with some of the initial biochemical experiments, Dr D. M. Brown and Professor A. R. Fersht for helpful discussions, and Dr S. E. V. Phillips for assistance with various aspects of the crystallographic comput-

ing. We acknowledge the financial support provided by an Usher Fellowship (to J.C.D.), an EMBO Short Term Fellowship (to R.S.B.), and NIH, NCI postdoctoral Fellowships 1 F22 CA02210-01 and 1 F32 CA06223-01 (to B.E.H.).

Received 17 January; accepted 19 April 1983.

- Britten, R. C. *r. Trav. Lab. Carlsberg* **32**, 371–380 (1962).
- Huff, J. W., Sastry, K. S., Gordon, M. P. & Wacker, W. E. C. *Biochemistry* **3**, 501–506 (1964).
- Farkas, W. R. *Biochim. biophys. Acta* **155**, 401–409 (1968).
- Farkas, W. R., Hewins, S. & Welch, J. W. *Chemico-Biol. Interactions* **5**, 191–200 (1972).
- Farkas, W. R. *Chemico-Biol. Interactions* **11**, 253–263 (1975).
- Werner, C., Krebs, B., Keith, G. & Dirheimer, G. *Biochim. biophys. Acta* **432**, 161–175 (1976).
- Jack, A., Ladner, J. E., Rhodes, D., Brown, R. S. & Klug, A. *J. molec. Biol.* **111**, 315–328 (1977).
- Hingerty, B. E., Brown, R. S. & Klug, A. *Biochim. biophys. Acta* **697**, 78–82 (1982).
- Jack, A., Ladner, J. E. & Klug, A. *J. molec. Biol.* **108**, 619–649 (1976).
- Hingerty, B., Brown, R. S. & Jack, A. *J. molec. Biol.* **124**, 523–534 (1978).
- Fersht, A. R. *Enzyme Structure and Mechanism*, 49–51 (Freeman, San Francisco, 1977).
- Brown, D. M. in *Basic Principles in Nucleic Acid Chemistry* Vol 2 (ed. Ts'o, P. O. P.) 1–90 (Academic, New York, 1974).
- Harrison, P. G., Healy, M. A. & Steel, A. T. *Inorg. chim. Acta* **67**, L15–L16 (1982).
- Usher, D. A. *Proc. natn. Acad. Sci. U.S.A.* **62**, 661–667 (1969).
- Roberts, G. C. K., Dennis, E. A., Meadows, D. H., Cohen, J. S. & Jardetzky, O. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1151–1158 (1969).
- Usher, D. A., Richardson, D. I. & Eckstein, F. *Nature* **228**, 663–665 (1970).
- Butzow, J. J. & Eichhorn, G. L. *Biopolymers* **3**, 95–107 (1965).
- Eichhorn, G. L., Butzow, J. J., Clark, P. & Tarien, E. *Biopolymers* **5**, 283–296 (1967).
- Butzow, J. J. & Eichhorn, G. L. *Biochemistry* **10**, 2019–2027 (1971).
- Burgess, J. *Metal Ions in Solution*, 264–267 (Ellis Horwood, Chichester, 1978).
- Harrison, P. G. *Co-ord. Chem. Rev.* **20**, 1–36 (1976).
- Kruger, K. *et al. Cell* **31**, 147–157 (1982).
- Zaug, A. J., Grabowski, P. J. & Cech, T. R. *Nature* **301**, 578–583 (1983).

Evidence on human origins from haemoglobins of African apes

Morris Goodman*, Gerhard Braunitzer†, Anton Stangl† & Barbara Schrank†

* Department of Anatomy, Wayne State University School of Medicine, Detroit, Michigan 48201, USA

† Max-Planck Institut für Biochemie, Abteilung Proteinchemie, D-8033 Martinsried bei München, FRG

Molecular data have influenced views concerning human origins, first, by supporting the genealogical classification of *Pan* (chimpanzee) and *Gorilla* with *Homo* rather than with *Pongo* (orang-utan)^{1,2} and, second, by suggesting that only a few million years separate humans and chimpanzees from their last common ancestor^{3,4}. Indeed, the cladistic distances in phylogenetic trees constructed from amino acid sequence data, on detecting many superimposed mutations, yielded a 'molecular-clock' divergence date between *Homo* and *Pan* of only 1–1.5 Myr BP⁵. This date, which is even more recent than that (4.2–

5.3 Myr BP)⁶ calculated using phenetic distances from immunological and DNA-hybridization comparisons (Table 1), is too near the present considering the existence of 3–4 Myr-old fossils of bipedal human ancestors⁷ (and a 5.5 Myr-old jaw fragment assigned to *Australopithecus*⁸). Perhaps decelerated sequence evolution occurred; alternatively, hominoid distances could have been underestimated, because chimpanzee and gorilla were represented mostly by sequences inferred from peptide amino acid compositions, as was the case for their haemoglobins^{9,10}. To help rectify this situation we report here the rigorously determined α - and β -haemoglobin amino acid sequences not only of chimpanzee (*Pan troglodytes*) and *Gorilla gorilla* but also pygmy chimpanzee (*Pan paniscus*). Our findings favour the explanation of decelerated evolution and point to selection preserving perfected haemoglobin molecules.

The pygmy chimpanzee sequence is shown in Fig. 1, with supporting data in Fig. 2. There is no difference in this sequence from that of either chimpanzee or man, and only two differences from gorilla. At position α 23, gorilla has aspartic acid instead of glutamic acid and at β 104, lysine instead of arginine. When these sequences are used with other known haemoglobin sequences in phylogenetic reconstructions by the maximum parsimony method, evidence is provided for cladistically joining *Pan* and *Gorilla* to *Homo* in Homininae rather than to *Pongo* in Ponginae, these two subfamilies being sister groups (Fig. 3a). Specifically, substitutions at β 87 (lysine $\rightarrow$ threonine) and β 125 (glutamine $\rightarrow$ proline) group *Pan*, *Homo* and *Gorilla* into Homininae. When all relevant amino acid sequence data are considered, breaking up the African ape–human clade with orang-utan adds at least 10 nucleotide replacements (NRs) over the most parsimonious score (2 NRs apiece for fibrinopeptides A and B, myoglobin and β -haemoglobin, and 4 NRs for carbonic anhydrase¹¹). Further parsimony evidence for a monophyletic chimpanzee–human–gorilla clade has been obtained from mitochondrial DNA nucleotide sequence data, representing close to 900 aligned positions of human¹², chimpanzee¹³, gorilla¹³, orang-utan¹³, gibbon¹³, mouse¹⁴ and ox¹⁵. Taking mouse and ox as outgroups of Hominoidea and calculating NR lengths for each possible dichotomous branching order among the five hominoid lineages, it was found¹⁶ that breaking up the human–chimpanzee–gorilla clade adds at least 19 NRs to the most parsimonious score.

Aside from supporting inclusion of *Pan* and *Gorilla* in Homininae, the parsimony evidence from amino acid sequence data hints at the possibility that *Pan* is more closely related to *Homo* than to *Gorilla*. One less NR is required with α -haemoglobin sequences by grouping *Pan* and *Homo* first before adding gorilla (Fig. 3a). This is due to aspartic acid at position α 23 in New World monkey, gibbon, orang-utan and gorilla chains as

Table 1 Comparison of 'clock' dates from different sets of molecular data

Ancestral node	Immunological		DNA hybridization (Myr BP)	Amino acid sequence (Myr BP)	
	Albumin (Myr BP)	Transferrin (Myr BP)			
Theria	125	X	X	117	200
Eutheria	90–100	X	X	90	154
Primates	73.5	63	91	51	87.1
Anthropoidea	35	35	35	20.5	35
Catarrhini	20.3	18.6	21.4	13.4	22.9
<i>Homo</i> – <i>Pan</i> –(<i>Gorilla</i>)	4.2(4.2)	4.6(4.6)	5.3(5.3)	1.3(1.8)	2.2(3.1)

Times for the Theria (Metatheria–Eutheria) and Eutheria ancestral nodes for albumin immunological distances are from Sarich and Cronin⁶ as described in the text of their article; the other divergence times from albumin and transferrin immunological distances and from DNA hybridization distances are calculated from Table 2 of their article using 35 Myr BP⁶ for the Anthropoidea (platyrrhine–catarrhine) ancestral node as the setting of the 'clock'. The divergence clock dates shown in the last two columns are from Table 9 of ref. 5; these dates are based on calculations using phylogenetic trees constructed from amino acid sequence data on up to 10 polypeptide chains (α - and β -haemoglobins, myoglobin, lens α -crystallin A, cytochrome c, fibrinopeptides A and B, and carbonic anhydrases I, II and III). In the second to last column, the setting of the clock is 90 Myr BP for Eutheria ancestral node. In the last column, 35 Myr BP for Anthropoidea ancestral node is used to set the clock, thus allowing more direct comparison with values from Sarich and Cronin⁶. Note that the so-called molecular clocks are not in markedly good agreement with one another. The value underlined for each set of clock dates is the setting for these calculations. 'X' Means that no clock date could be calculated because no corresponding distance value was given in Table 2 of Sarich and Cronin⁶.



Fig. 1 Complete amino acid sequence of pygmy chimpanzee haemoglobin. The sequence is deduced in the liquid phase sequenator using the Quadrol program²⁶ and the *N,N*-diethylaminopropyne program²⁷. The amino acid analysis of the most important peptides is given in Fig. 2. The amino acid sequences of human haemoglobin A and of the haemoglobins of pygmy chimpanzee and chimpanzee are all identical.

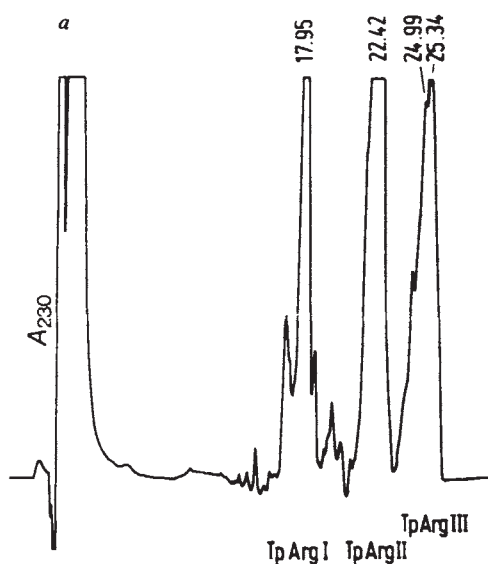


Fig. 2 *a*, Amino acid analysis of most important peptides for sequencing pygmy chimpanzee haemoglobin. Tp Arg, tryptic peptides after blocking lysine side chains. Pro Pep., C-terminal peptide after hydrophilic splitting of Asp-Pro bond²⁸. The separation of most peptides was performed by HPLC chromatography. Separation of the 'Arginine' peptides from the α -chains by HPLC (apparatus: Beckman Instruments, RP 2 column acetonitrile gradient; ammonium acetate buffer pH 6.0).

b

BOROBO B-CHAINS

	TP1 Arg	TP2 Arg	TP3 Arg		Pre Pop.	Chain Analysis
Pos.	1-31	32-92	93-141		99-141	
Lys	3.00	5.24	3.25		3.17	11.1
His	1.21	6.27	3.20		3.13	9.00
Arg	1.11	1.08	1.08		1.05	3.00
Asp	2.16	6.80	3.07		2.40	12.32
Thr	1.08	3.83	3.70		3.80	8.74
Ser	1.08	4.92	4.80		4.80	10.80
Glu	2.90	1.15	1.08		1.27	5.30
Pro	1.01	2.94	2.86		2.86	7.95
Gly	3.84	3.19				7.30
Ala	7.06	8.18	5.85		6.26	20.78
Cys			0.70		0.90	0.90
Val	3.05	4.15	5.90		4.80	12.70
Met		1.85				1.70
Leu	2.20	7.12	9.10		8.86	18.20
Tyr	1.01	1.08	0.95		1.05	2.90
Phe		4.22	3.12		3.13	7.06
Try	0.90					0.80
Total	31	61	49		47	141

BOROBO B-CHAINS

	TP1 Arg	TP2 Arg	TP3 Arg	TP4 Arg		CR1	CR2	Pre Pop.	Chain Analysis
Pos.	1-30	31-40	41-74	75-146		1-55	56-146	70-146	
Lys	2.14		6.29	2.92		2.26	8.84	3.14	11.40
His	1.07		4.36	3.62		1.21	6.10	3.97	9.10
Arg	1.07	1.07	1.01			2.08	1.10	0.99	3.00
Asp	2.09		8.70	2.19		4.08	8.80	3.14	12.80
Thr	1.92	1.06	2.80	1.14		3.91	3.07	1.07	6.88
Ser	1.01		3.80			2.87	2.88		4.78
Glu	3.96	1.11	3.20	3.13		5.99	4.62	3.97	11.30
Pro	0.96	1.02	2.80	2.18		2.87	3.80	2.70	7.40
Gly	3.85		5.90	3.18		4.78	8.11	3.30	12.60
Ala	3.00		5.20	7.00		4.17	10.90	6.94	14.75
Cys			0.80	0.90			1.80	1.00	1.80
Val	4.80	1.75	4.30	6.70		7.04	10.70	7.18	17.80
Met			0.91						0.81
Leu	3.05	1.98	8.20	5.06		6.20	11.90	8.29	18.30
Tyr		1.03		1.98		1.04	1.96	1.98	3.10
Phe			5.89	2.14		3.21	5.18	2.97	8.18
Try	1.00	0.90				1.80			1.70
H-Ser						1.00			
Total	30	10	64	42		85	91	47	146

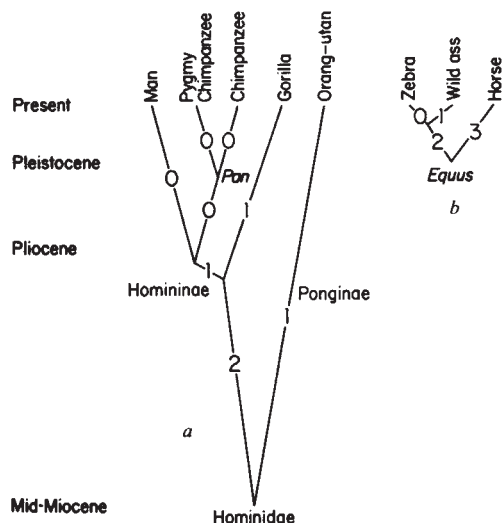


Fig. 3 Hominid (a) and equine (b) radiations as deduced from maximum parsimony trees of α - and β -haemoglobin sequences. Numbers of NRs are shown on the branches. They result from the following substitutions recorded in the single-letter code for amino acids²¹. a, Between Hominidae ancestral node and orang-utan: α 12 A $\rightarrow$ T; between Hominidae and Homininae ancestral nodes: β 87 K $\rightarrow$ T, β 125 Q $\rightarrow$ P; between Homininae ancestral node and gorilla: β 104 R $\rightarrow$ K; between Homininae and *Pan-Homo* ancestral nodes: α 23 D $\rightarrow$ E. b, Between *Equus* ancestral node and horse: β 52 A $\rightarrow$ G, β 87 Q $\rightarrow$ A; between *Equus* and zebra-wild ass ancestral nodes: α 20 H $\rightarrow$ N, α 131 S $\rightarrow$ T; between zebra-wild ass ancestral node and wild ass: α 23 E $\rightarrow$ D. Immunological evidence for cladistically grouping pygmy chimpanzee with chimpanzee in the genus *Pan*²⁹ is now supported by amino acid sequence results on carbonic anhydrase I (D. Hewett-Emmett, unpublished data). The positioning of the divergence node for pygmy chimpanzee and chimpanzee reflects views³⁰ based on immunological and electrophoretic distance data.

compared with glutamic acid at this position in human, chimpanzee and pygmy chimpanzee chains. Amino acid sequences of γ -haemoglobin chains also point to a *Homo-Pan* clade distinct from gorilla. There are two non-allelic chains in the hominines, $^{\alpha}\gamma$ characterized by glycine at position 136 and $^{\beta}\gamma$ with alanine at this position. The two sequences in chimpanzee as inferred from amino acid composition data are identical to their human counterparts¹⁷. In contrast, gorilla $^{\alpha}\gamma$ (as deduced from nucleotide sequence data obtained on gorilla γ -haemoglobin genes)¹⁸ has arginine at position 104 whereas human and chimpanzee $^{\alpha}\gamma$ chains have lysine. The presence of arginine at position 104 in Old World monkey (*Macaca* and *Papio*) γ -haemoglobin chains¹⁹ raises the possibility that human and chimpanzee $^{\alpha}\gamma$ have the derived rather than primitive residue.

With mitochondrial DNA sequences there are two trees of lowest NR length grouping *Pan* either with *Gorilla* or with *Homo*¹⁶; phenetically, however, *Pan* shows more base matches with *Homo* than with *Gorilla*^{13,16} as it also does with nuclear DNA sequences²⁰. Clearly it will be difficult to resolve the *Homo-Pan-Gorilla* trichotomy into two branching events if, as may be suspected, the second branching occurred shortly after the first.

That proteins such as globins show a strong—although uneven—tendency to increasing sequence divergencies with elapsed time is evident in both amino acid difference matrices^{21,22} and phylogenetic reconstructions⁵. Accepting empirical observations that this tendency operates in a majority of clades, our finding of almost no sequence differences among hominines in either their α - or β -haemoglobin chains argues for a relatively late ancestral separation of *Homo*, *Pan* and *Gorilla*. Nevertheless, even if common ancestry existed until 5–6 Myr BP (just before the appearance of *Australopithecus* fossils), the rate of hominine haemoglobin evolution was slow, averaging about only 6 NRs per 100 codons per 10^8 yr (that

is, 6 NR%), in contrast to much faster rates averaging about 29 NR% for β -haemoglobin sequences and 25 NR% for α -sequences during eutherian phylogeny⁵.

To highlight the slowness of haemoglobin evolution during the hominine radiation, we can compare hominine with equine rates, taking advantage of the fact that α - and β -haemoglobin sequences have been determined for horse (*Equus caballus*), zebra (*Equus zebra*) and wild ass (*Equus hemionus*)²³. Preceded by *Pliohippus*, the genus *Equus* does not appear in the fossil record until the Pleistocene²⁴. Yet, during the short span of time (2 Myr BP to present) of this equine radiation, 2–3 NRs accumulated per equine lineage for the two haemoglobin chain types (Fig. 3b), compared with no changes in *Homo* and *Pan*. Even by the conservative assumption that *Homo* and *Pan* had no older separation than zebra and horse, and given a Poisson distribution with a mean of 5 NRs based on *Equus*, the probability of obtaining zero NRs as for *Homo-Pan* is less than 0.007—thus eliminating the hypothesis of homogeneous rates.

The slow rate of globin sequence evolution in hominines indicates prior selection in the pre-hominines of haemoglobin molecules with finely tuned, perfected adaptations. As previously reported⁵, we find that between eutherian and hominine ancestral nodes, the fastest evolving positions are $\alpha_1\beta_1$ contact sites, important functional positions which show few normal human variants²⁵. Also mutations to proline were fixed at exterior helical positions α A2 and β A2 and at $\alpha_1\beta_1$ contact positions β D2 and β H3. These additional prolines, all at the beginning of helices, by producing more stably directed helices, presumably improved subtler functions in the haemoglobin molecule. Furthermore, almost none of the many seemingly normal haemoglobin variants²⁵ are at polymorphic frequencies, possibly because even these variants harm the finer adaptations of human haemoglobin. The slow rate of haemoglobin evolution in African apes and humans suggests that organismal evolution in our earlier ancestors may have depended on amino acid sequence changes in proteins as well as regulatory changes in gene expression.

We thank Drs A. Spiegel, R. Faust and H. Wiesner for the pygmy chimpanzee, chimpanzee and gorilla blood samples, also Dr Gerard Joswiak and John Czelusniak for valuable discussion of the statistical significance of our findings. Phylogenetic analysis of the sequence data was supported by NSF grant DEB 78-10717.

Received 30 December 1982; accepted 7 April 1983.

- Goodman, M. *Ann. N. Y. Acad. Sci.* **102**, 219–234 (1962).
- Andrews, P. & Cronin, J. E. *Nature* **297**, 541–546 (1982).
- Sarich, V. M. & Wilson, A. C. *Science* **158**, 1200–1203 (1967).
- Goodman, M. *Hum. Biol.* **54**, 247–264 (1982).
- Goodman, M. *Prog. Biophys. molec. Biol.* **38**, 105–164 (1981).
- Sarich, V. M. & Cronin, J. E. in *Molecular Anthropology* (eds Goodman, M. & Tashian, R. E.) 141–170 (Plenum, New York, 1976).
- Johanson, D. L. & White, T. D. *Science* **203**, 321–330 (1979).
- Patterson, B., Behrensmeier, A. K. & Sill, W. *Nature* **226**, 918–921 (1970).
- Rifkin, D. B. & Konigsberg, W. *Biochim. biophys. Acta* **104**, 457–461 (1965).
- Zuckerkandl, E. & Schroeder, W. A. *Nature* **192**, 984–985 (1961).
- Tashian, R. E., Hewett-Emmett, D. & Goodman, M. in *Isozymes* Vol. 7 (eds Rattozzi, M. C., Scandalias, J. G., Siciliano, M. T. & Whitt, G. S.) 79–100 (Liss, New York, 1983).
- Anderson, S. *et al. Nature* **290**, 457–465 (1981).
- Brown, W. M., Prager, E. M., Wong, A. & Wilson, A. C. *J. molec. Evol.* **18**, 225–239 (1982).
- Anderson, S. *et al. J. molec. Biol.* **156**, 683–717 (1982).
- Bibb, M. H. *et al. Cell* **26**, 167–180 (1981).
- Goodman, M., Olson, C. B., Beeber, J. E. & Czelusniak, J. *Acta zool. fenn.* **169**, 19–35 (1982).
- DeJong, W. W. *Biochim. biophys. Acta* **251**, 217–226 (1971).
- Scott, A. F. *et al. Am. J. hum. Genet.* **34**, 193A (1982).
- Mahoney, W. C. & Nute, P. E. *Biochemistry* **19**, 4436–4442 (1980).
- Hoyer, B. H., Van de Velde, N. W., Goodman, M. & Roberts, R. B. *J. hum. Evol.* **1**, 645–649 (1972).
- Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5 (National Biochemical Research Foundation, Silver Springs, Maryland, 1972).
- Dickerson, E. R. & Geis, L. *Hemoglobin: Structure, Function, Evolution, and Pathology* (Benjamin/Cummings, Menlo Park, 1983).
- Mazur, G. & Braunitzer, G. *Hoppe-Seyler's Z. physiol. Chem.* **363**, 59–71 (1982).
- Romer, A. S. *Vertebrate Paleontology* (University of Chicago Press, 1966).
- Bunn, H. F., Forget, B. G. & Ranney, H. M. *Human Hemoglobins* (Saunders, Philadelphia, 1977).
- Edman, P. & Begg, G. *Eur. J. Biochem.* **1**, 80–91 (1967).
- Braunitzer, G. & Schrank, B. *Hoppe-Seyler's Z. physiol. Chem.* **351**, 417–418 (1970).
- Jauregui-Adell, J. & Marti, J. *J. analyt. Biochem.* **69**, 468–473 (1975).
- Goodman, M. *Hum. Biol.* **35**, 377–436 (1963).
- Zihlman, A. L., Cronin, J. E., Cramer, D. L. & Sarich, V. M. *Nature* **275**, 744–746 (1978).

Overlapping spreading centres on East Pacific Rise

FROM a detailed Seabeam investigation of the East Pacific Rise¹ (EPR), which provided continuous along-strike coverage of the rise axis neovolcanic zone between 8°N and 18°N, Macdonald and Fox¹ claim the discovery of a new kind of volcano-tectonic geometry associated with fast spreading centres. This geometry—overlapping spreading centres (OSC)—has been inferred from discontinuities in the morphology of the neovolcanic zone at several locations along the rise axis. The discontinuities are characterized by two overlapping neovolcanic zones enclosing an elongate depression. The existence of similar features on the southern EPR between 20°S and 7°N has been reported by Lonsdale².

Macdonald and Fox¹ suggest that the inferred OSCs represent an accretion geometry never before observed, and in particular, that have not been detected during detailed studies of the slow spreading Mid-Atlantic Ridge (MAR). Lacking Seabeam bathymetric information beyond the neovolcanic zone, the authors use wax model studies to suggest that overlap-

ping spreading centres represent a transient phenomenon of lithospheric accretion qualitatively similar to propagating rifts.

We have placed their OSCs within the regional tectonic framework by locating them on the bathymetric-tectonic map of the EPR from Klitgord and Mammerickx³ (Fig. 1). The heavy dashed lines denote the location of fracture zones (FZ) predicted by Klitgord and Mammerickx³ from bathymetry and marine magnetic anomalies whereas the light dashed lines are additional FZs inferred by us from discontinuities in the bathymetric contours. Four of the OSCs (B, D, F and H of Fig. 1) are within 5 km of the predicted fracture zones which intersect the central magnetic anomaly. Two of the remaining four OSCs (A and G, Fig. 1) are associated with distinctive discontinuities in the bathymetry which may represent FZs on oceanic crust older than anomaly 2'. OSCs C and E (Fig. 1) may correspond to the older FZs mapped by Klitgord and Mammerickx but the existing bathymetric data are insufficient to make such a tie reliably. The small distance between some of the FZs and OSCs can be attributed to the finite width of the OSCs and sparse nature of the data used to locate the FZs.

This correlation between FZs and OSCs indicates that Macdonald and Fox have recorded the morphological expression of 'zero offset' (or small offset) transform zones^{4,5} which are the active parents of zero-offset fracture zones similar to the ones reported in the North Atlantic^{4,6}, and that the OSC-transform zones have been stable features on the EPR accreting boundary for at least 5 Myr.

The existence of overlapping or *en echelon* spreading centres has been postulated previously by Ramberg *et al.*⁷ from the multibeam echosounder investigation of the Mid-Atlantic FAMOUS area⁸. Overlapping rift zones have been documented on Iceland⁹. In the FAMOUS area, the 20-km offsets between the neovolcanic zones are larger than on the EPR and each zone terminates in a nodal depression located in the active transform zone. Similar nodal depressions exist at intersections between spreading centres and large offset transform zones in North Atlantic and Pacific^{10,11}. Minor (<10 km) offset transform zones in the central North Atlantic TAG area^{12,14} exhibit a single nodal depression resembling the OSC morphology observed by Macdonald and Fox¹. A geometry of finite-width transform zones between spreading centres, proposed by Schouten *et al.*⁵ to explain asymmetric formation of sea floor at MAR spreading centre/transform-fault intersections, requires that both spreading centres be active (overlap) in the transform zone.

The continuity and stability of small-to zero-offset transform zones has been inferred in the North Atlantic using local surveys of sufficient data density to map small discontinuities in the magnetic lineation pattern^{4,6}. During its 175 Myr of evolution, the central North Atlantic accreting boundary consisted of a more or less stable string of closely spaced spreading centres with intervening transform zones about every 50 km (ref. 6); the offsets on these transform zones have varied through time as the accreting boundary has responded to changes in plate motion. The stability of zero-offset transform zones is probably a thermal regime stability associated with finite-length magma chambers^{4,6}. The data set of Macdonald and Fox¹ does not substantiate whether OSCs propagate along the EPR accreting boundary but the correlation with the FZs (Fig. 1) indicates that they probably do not migrate. The fact that many OSCs now are recognized on the EPR^{1,2} suggests that their morphological expression, consisting of two overlapping neovolcanic zones enclosing an elongate depression, is stable rather than transient.

Contrary to the conclusions reached by Macdonald and Fox¹ that the OSCs are transient features confined to fast spreading regimes, their observed OSC geometries confirm the present-day

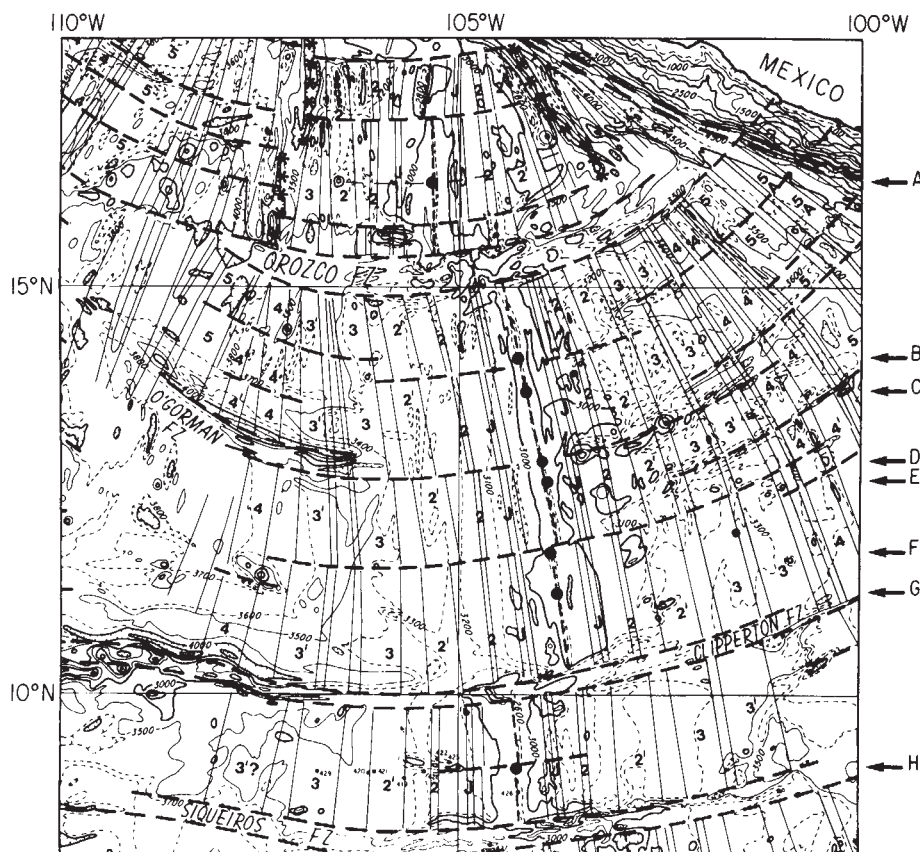


Fig. 1 Bathymetric-tectonic map of the East Pacific Rise with predicted fracture zone traces (dashed lines) and marine magnetic lineations (after ref. 3), and with location of overlapping spreading centres (OSCs, ● after ref. 1). The correlation between OSC locations and predicted fracture zones indicates that these OSCs represent zero-offset transform zones^{4,5} that have been stable features on the East Pacific accreting boundary for at least 5 Myr (see text for details).

existence of stable small- to zero-offset transform zones on the fast-spreading EPR accreting boundary that were predicted by Klitgord and Mammerickx³ and that are similar to the ones reported from the slow-spreading North Atlantic^{4-7,9}. Their observation provides further evidence for a continued separation between individual spreading centre cells when no lateral offset exists between the cells^{4,6}. The average spacing between spreading centre cells in Pacific and Atlantic Ocean seems to be similar despite a 4:1 ratio in spreading rates. This predicts an orderly pattern of sea floor in both oceans with a temporally stable string of spreading centre cells and intervening transform zones leaving its trace parallel to the flowlines of relative plate motion in the form of narrow ribbons of normal oceanic crust and intervening fracture zones.

We thank Mike Purdy, Bill Bryan, Anne Trehu and Debbie Hutchinson for constructive criticism. We acknowledge support from the Office of Naval Research and the NSF. Pam Foster and Maureen Carragher processed the manuscript. Woods Hole Oceanographic Institution contribution no. 5312.

HANS SCHOUTEN*
KIM D. KLITGORD†

* Woods Hole Oceanographic
Institution and

† US Geological Survey,
Woods Hole,
Massachusetts 02543, USA

1. Macdonald, K. C. & Fox, P. J. *Nature* **302**, 55-58 (1983).
2. Lonsdale, P. *EOS* **63**, 1108 (1982).
3. Klitgord, K. D. & Mammerickx, J. *J. geophys. Res.* **87**, 6725-6750 (1982).
4. Schouten, H. & White, R. *Geology* **8**, 175-179 (1980).
5. Schouten, H., Karson, J. & Dick, H. *Nature* **288**, 470-473 (1980).
6. Schouten, H. & Klitgord, K. *Earth planet. Sci. Lett.* **59**, 255-266 (1982).
7. Ramberg, I. B., Gray, D. F. & Reynolds, R. G. M. *Bull. geol. Soc. Am.* **88**, 609-620 (1977).
8. Phillips, J. D. & Fleming, H. S. *Map and Chart Ser. MC-19* (Geological Society of America, 1978).
9. McMaster, R. L., Schilling, J.-G. & Pinet, P. R. *Nature* **269**, 663-668 (1977).
10. Sleep, N. H. & Biehler, S. J. *geophys. Res.* **75**, 2748-2752 (1970).
11. Fox, P. J. *et al. EOS* **62**, 1049 (1981).
12. Rona, P. A. *et al. Bull. geol. Soc. Am.* **87**, 661-674 (1976).
13. Rona, P. A. & Gray, D. F. *Bull. geol. Soc. Am.* **91**, 485-494 (1980).

MACDONALD REPLIES—Schouten and Klitgord suggest that several of our OSCs coincide with the locations of real and hypothetical fracture zones presented in Klitgord and Mammerickx's¹ synthesis of bathymetric and magnetic data for the northern EPR (Fig. 1). If true, this implies that OSCs are stable for tens of millions of years. Based on wax models of spreading centre processes, however, we suggest that OSCs represent a rapidly evolving expression of how oceanic lithosphere is accreted at fast spreading rates². In the development and evolution of OSCs actually observed in our wax models, the process begins and ends with a long-

wavelength sinuous bend in the trend of the spreading centre (see Fig. 3 of ref. 2). We hypothesized that epixodes of volcanism along the neovolcanic zone propagate or migrate along strike in some episodic way, and often fail to meet at places where there is a sinuous bend in the rise axis. The neovolcanic zones proceed to overlap, a short-lived shear zone develops between them, and eventually one of the OSCs prevails as the other is abandoned. The prevailing OSC is left with a bend along its strike.

It is possible that OSCs achieve some degree of stability from this relict bend, because subsequent episodes of volcanism may preferentially overlap each other once again at or near this bend, that is, the same place as in the previous cycles. Thus, while any instantaneous OSC system is unstable, rapidly evolving, and characterized by non-rigid plate behaviour, the process may recur near the same place along the spreading axis. If this is true, OSCs will leave traces in older lithosphere which look like fracture zones, but which have a fine-scale structure of volcano-tectonic trends which are 5°-20° oblique to the rise axis, and 70°-85° oblique to expected fracture zone trends. Also, the OSC scars left in older lithosphere may not trace clear small circles about the pole of opening for the plates. The non-rigid behaviour of OSCs would render them less stable spatially than rigid-plate transform faults, and some along-strike migration is likely to occur. The OSC scars should appear as groups of short cusped ridges which disrupt the otherwise linear abyssal hill terrain. Schouten and Klitgord's observations can be reconciled with the evolutionary model for OSCs which we have proposed, and definitive tests can be made by conducting high resolution observations in older lithosphere. Until such observations are made, however, we remain sceptical and do not rule out the possibility that the proposed correlation between our OSCs and major fracture zones is fortuitous.

Note that four out of eight OSCs do not correspond to presently active transform faults on the Klitgord and Mammerickx chart (Fig. 1A, C, D, G) and that two of their hypothetical transform faults in young crust do not show any offset or overlap in our high resolution seabeam charts of the EPR. The fracture zone traces are not controlled by magnetic data as they are in the Atlantic, but rely primarily on widely spaced bathymetric profiles^{3,4}. Thus many of the fracture zone locations and identifications in Fig. 1 are subject to considerable uncertainty (J. Mammerickx, personal communication).

At present, there is insufficient high resolution data to test whether (1) OSCs are totally stable, tracing small circles about the pole of opening (as Schouten and Klitgord suggest), (2) OSCs have some degree of spatial stability but are

not likely to produce clear, small circle scars about the pole of opening (as we outline here), or (3) OSCs are entirely migratory as well as transient, producing randomly located cusped ridges in the abyssal hill terrain. A high resolution field programme in the Pacific using Seabeam and Seamarc is presently under way which should provide data to test these three hypotheses.

Finally, our experience with OSCs in the Pacific versus zones of apparent spreading centre overlap in the Atlantic, such as Fracture Zone B,^{5,6} suggests that the tectonic processes may be quite different. The slow-spreading Atlantic is characterized by distinct lateral shifts of true transform faults with minor instantaneous overlapping of the adjacent offset spreading centres. The fast-spreading Pacific is characterized by large spreading centre overlaps with no evidence for transform fault structures in the overlap zone. Thus OSCs have a different volcano-tectonic expression than Schouten's "zero-offset" transform faults³.

Note added in proof: Detailed Seabeam and Seawave studies near the OSCs at 12° 54' N and 11° 45' N (D, F in Fig. 1) document that the faulted abyssal hill terrain is uninterrupted adjacent to the OSCs, and that these two OSCs have not left fracture zone traces as suggested in Fig. 1. Instead, cusped ridges and depressions were mapped in older lithosphere as predicted by our model, consistent with options (2) or (3) above, but not with (1).

KEN C. MACDONALD
Department of Geological Sciences,
University of California,
Santa Barbara,
California 93106, USA

1. Klitgord, K. D. & Mammerickx, J. *J. geophys. Res.* **87**, 6725-6750 (1982).
2. Macdonald, K. C. & Fox, P. J. *Nature* **302**, 55-58 (1983).
3. Schouten, H. & White, R. *Geology* **8**, 175-179 (1980).
4. Schouten, H. & Klitgord, K. *Earth planet. Sci. Lett.* **59**, 255-266 (1982).
5. Macdonald, K. C. & Luyendyk, B. P. *Bull. geol. Soc. Am.* **88**, 621-636 (1977).
6. Reid, I. & Macdonald, K. C. *Nature* **246**, 88-90 (1973).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

Across the chessboard of life

David L. Hull

The Secular Ark: Studies in the History of Biogeography.

By Janet Browne.

Yale University Press: 1983. Pp.265. \$27.50, £21.

IN *The Secular Ark*, Janet Browne compares the surface of the Earth to an immense chessboard on which species are distributed. From the beginning the task of the biogeographer was to discern this pattern and to postulate processes that might have produced it.

The Bible provided the K-P4 in this game. According to biblical literalists, God created all living creatures in the Garden of Eden. From there they spread across the face of the land only to be annihilated by the Great Flood. After Noah's Ark settled on Mount Ararat, the process was repeated. Although no traces remained of the first great dispersal, early biogeographers were presented with the problem of how giraffes got from Ararat to Africa, armadillos to South America and the partridge to Scotland. As numerous voyages of discovery increased the number of known species, however, the Ark began to strain at its seams. Just as Newtonian physicists relegated problems surrounding action-at-a-distance to the closet, biogeographers simply ignored this problem once it became too intractable and concentrated on the flood itself. For example, Linnaeus never mentions the Ark in his description of the origin of species on a mountainous equatorial island that somehow incorporated all possible habitats. As the ocean surrounding this island receded, organisms spread out from Eden. (Browne relegates Linnaeus's later, more sophisticated views to a single parenthetical aside.)

Browne proceeds rapidly from such early, primitive theories to the more secular views of Alexander von Humboldt (1769–1859), Augustin de Candolle (1778–1841) and Robert Brown (1773–1858), giving particular attention to de Candolle's "botanical provinces" and Humboldt's "botanical arithmetic". The prevailing view of the time was that species existed in reasonably discrete, integrated natural assemblages or biotas. To aid in their discovery, Humboldt devised an elementary mathematical technique for determining proportions of various sorts of species; just as physicists constructed the periodic table by means of numerical ratios exhibited by the physical elements, Humboldt proposed to analyse the living world. Browne's failure to discuss Linnaeus's later views is especially unfortunate here, because he had already developed a system in which crosses between a few elemental species could generate all derivative species, much in the same way that compounds can be generated by various com-

binations of the elements.

Although the enthusiasm for Humboldt's botanical arithmetic was short-lived, it did fire the imagination of one budding geologist-naturalist — Charles Darwin (1809–1882). The theories of such geologists as Charles Lyell (1797–1875) forced biogeographers to become much more cognizant of time and process.

According to one theory popular at the time, a globally uniform climate allowed single primitive populations to live evenly distributed over the surface of the Earth. As climates became diverse, these single populations became increasingly fragmented. This view was soon supplanted by a belief in a separate centre of creation for each biological province. Lyell thought that species appeared piecemeal by some unknown but natural means, while Louis Agassiz (1807–1873) argued that they appeared all at once by divine fiat. The chief anomaly to be explained by these theories was the appearance of the same species in more than one province. Were these species created several times, once in each province, or might not species wander from province to province? Why was it that geographically neighbouring species tended to be so similar to each other? Together, geology and biogeography set the stage for Darwin's solution to the mystery of mysteries.

Because so much attention has been paid to the confrontation between biblical literalism and Darwin's theory, an equally important conflict has been all but ignored

— that between transcendental idealism and more empirical views of the world. At the time explanations in terms of "archetypes" and "ideal plans" were considered perfectly legitimate on the Continent. Showing that several species all exhibited the same *Bauplan* explained the similarities exhibited. This view of science was just beginning to get a good toe-hold in the British Isles through the work of Richard Owen (1804–1892), Edward Forbes (1815–1854) and that rising young talent, T.H. Huxley (1825–1895). Forbes, for example, explained the geological distribution of fossil species in terms of the action of the principle of polarity acting from opposite ends of an ideal sphere. Other nineteenth-century scientists in Great Britain, such as Lyell and Darwin, could make no sense of such "explanations". They failed to see how they explained anything. Explaining the similarities exhibited by species in terms of common ancestors, that was quite a different matter. As Browne argues, the Darwinian revolution was as much a triumph of a particular view of science as it was of a particular view of the origin of species.

As the reader can surely tell from this review, Browne has accomplished the impossible. She has written a scholarly yet entertaining history of one of the most fascinating series of episodes in science. As anyone who reads the pages of this journal also knows, the story is far from over as dispersalist biogeographers do battle with vicariance and panbiogeographers; even a latter-day version of transcendental idealism has cropped up again. The possibility of checkmate in this game continues to recede into the distant future as ineluctably as the waters of the Great Flood receded from Mount Ararat. □

David L. Hull is Distinguished Professor in the Department of Philosophy at the University of Wisconsin-Milwaukee, Wisconsin.

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

Two by two: a seventeenth-century interpretation of the loading of the Ark, from Athanasius Kircher's *Arca Noë* (1675).

Cognitive demons and mental lapses

John C. Marshall

Absent-Minded? The Psychology of Mental Lapses and Everyday Errors.

By James Reason and Klara Mycielska.
Prentice-Hall: 1982. Pp.262.
Pbk \$6.95, £5.90.

It is unreliably reported that the Reverend William A. Spooner after accidentally spilling some salt on the tablecloth calmly proceeded to "remove the stain" by pouring a little claret over it. Slips of the tongue, ear, hand and mind are ubiquitous in all our lives and are normally a source of constant delight and amusement — except when, for example, a depressed first officer mishears his captain's jovial "cheer up" as "gear up" and promptly retracts the undercarriage while the aircraft is still racing down the runway.

Although Reason and Mycielska do have a chapter on "catastrophic lapses" (from the Charge of the Light Brigade, through the sinking of *H.M.S. Victoria*, to the Tenerife runway collision) most of their examples are taken from everyday errors whose consequences are no worse than acute embarrassment. One good justification for this concern with seeming trivia is that Reason and Mycielska make out a compelling case that catastrophic errors are simply a subset of "ordinary" errors, "distinguished by circumstantial rather

than psychological considerations". The "Freudian slip", where the perpetrator is unconsciously motivated towards disaster, is a very rare beast indeed. Reason and Mycielska therefore argue that accident researchers would do well to consider "the full range of human errors, irrespective of their consequences". Even if this were not so, everyday lapses would still provide invaluable reminders of the complexity of the "automatic pilots" which steer the 99 per cent of our actions we do not (and should not) think about: it is only when we miss the last step that the subtlety of walking down stairs strikes us. And, as the intellectual millipede discovered, 'tis all for the best that most of our activities can be discharged fluently without our paying over much conscious attention to their control.

Reason and Mycielska accordingly distinguish errors of judgement and planning from cases where a satisfactory plan goes awry in execution. On the basis of an extensive questionnaire and diary study, they propose a preliminary taxonomy for the latter cases and establish some basic incidence statistics: "absent-minded" errors occur when habitual tasks that have acquired some degree of automaticity are performed in familiar environments; people seem to differ reliably in their proneness to minor slips and lapses, and this individual liability is not limited to particular domains (memory, action, recognition or language) but rather appears across the board; the occurrence of such errors is associated with distraction and preoccupation but not with obvious emotional or environmental stress, although liability to lapses is correlated with the vulnerability of an individual to stress.

There emerges an overall picture of the mind as a congeries of quasi-autonomous modules ("cognitive demons") whose operation is only intermittently controlled and checked by conscious mental activity. Such control requires focal attention, a strictly limited resource. Automatization of skilled action reduces the frequency with which it is necessary to draw upon attentional resources, at the cost of occasional error as a "demon" is either captured by an environmental variable extraneous to the plan being executed, or is pushed aside by a "demon" of greater intrinsic strength.

Readers familiar with the history of neuropsychology will recognize in this picture John Hughlings Jackson's politico-social (*Evolution and Dissolution of the Nervous System*, 1884) and Sigmund Freud's "economic" model of the mind (*Project for a Scientific Psychology*, 1895). We should now attempt to turn these metaphors into genuine theory with deductive depth and predictive power. ||

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford.

Nutrition: the great tradition

John Rivers

The Elements of the Science of Nutrition.

Reprint of the fourth edition,
first published in 1928.

By Graham Lusk.

Academic: 1983. Pp.844. \$38.50, £26.50.

THERE have been two great periods in the science of nutrition, and Graham Lusk's book reflects both of them. Nutrition began as a physiological science, the nineteenth-century investigation of intermediary metabolism by the only means available, the comparison of input and output. In the early years such studies were both confused and confusing, but as the laws of the conservation of energy became clearly established a body of scientific knowledge was assembled. Developments were shaped by two fortuitous coincidences. Most of the nitrogen in food was in the α -amino groups of protein which provided a natural label for studying protein metabolism, while energy exchange, and, to a first approximation, the nature of the nutrients oxidized, could be predicted from respiratory exchange. Thus in the nineteenth century there was a flowering of work on protein and energy metabolism in the whole organism which is associated with such great names as Prout, Pereira, Leibig, Rubner and Voit.

By the end of the century this approach was wearing thin. It must have seemed intellectually bankrupt; a spirit of reductionism was in the air. Claude Bernard captured it when he snarled derisively: "We may of course strike a balance between what a living organism takes in as nourishment and what it gives out in excretion; but the results would be mere statistics incapable of throwing light on the inmost phenomena of nutrition in living beings". The discovery of the vitamins was an early triumph for Bernard's more molecular approach, and in this second golden age can be seen the schism (at once clear and yet difficult to define) between the biochemical and purely nutritional investigations of nutrient metabolism.

Graham Lusk's *The Elements of the Science of Nutrition*, its title redolent of Euclid's *Elements*, is a product of the first golden age. It is by no means the end of that great tradition, but it does represent its greatest synthesis. But the book was an anachronism, obsolescent when it first appeared in 1906. This reprint is Lusk's own copy of the most famous fourth edition of 1928 which must have stood like a dinosaur out of its time. The metabolism of protein and of energy take up most of its 850 pages — vitamins are discussed only in one brief chapter on the nutritive value of foods which makes up less than five per cent of the book. In this it is the obverse of

Journals' review issue 1983

On October 6th *Nature* will publish the third review supplement devoted to science journals. The two previous supplements (covering journals first published between January 1978 and May 1980, and June 1980 and May 1981) appeared in *Nature* 293, 341–369 (1981) and 299, 491–514 (1982).

Criteria for inclusion of a journal in the 1983 issue are that:

- the first number appeared, or the journal was re-titled, between June 1981 and May 1982;
- it is published at least three times a year;
- the main language used is English.

Broadly, periodicals of professional interest to scientists will be considered for review, with the exception of abstracts' journals. Journals appearing prior to June 1981 but after January 1978, and not covered in the previous issues, will also be considered.

Publishers and societies are invited to submit four sample issues of journals satisfying the above criteria, including the first and most recent numbers, to the Review Editor, *Nature*, 4 Little Essex St, London WC2R 3LF, England (London 836 6633 ext 2570) as soon as possible.

most texts with which it was contemporary, texts where the new and exciting evidence about vitamins clearly fascinated the writers and the macronutrients were somehow *passé*.

Lusk, a pupil of Voit's and a proud linear descendant of the nineteenth-century school of nutritional physiology, regarded the essence of nutritional science as a holistic view of metabolism. He was certainly out of step with his time and reading this magnificent book one is taken by the thought that maybe he didn't care. Certainly, now at a time when preoccupation with vitamin deficiencies has by and large receded from the headlines, what Lusk wrote emerges as vital and alive. We are today gripped by the problems which fascinated him. How precisely can the requirements for protein and energy be defined and how are they related? How far is a reduction in energy expenditure the cause of obesity? What are the limits of adaptation to low food intakes?

But Lusk's current appeal does not come simply from the now-fashionable nature of his interests. What he has to say, the data he so painstakingly presents, are also of current importance. Science books usually date very rapidly and become of interest only to the historian not the scientist. Lusk's detailed exposition of his own experimental work, however, and of the now almost totally unread works of his German teachers, is a rich vein yet to be mined. For example, it is commonly assumed that whole-body protein and energy equilibrium are established in the adult at the same energy intake, a view specifically espoused in the current WHO/FAO Protein and Energy Requirements report and implicit in much that is written about requirements. Lusk will have none of it, presenting careful (and still indisputable) experimental results to support his view that "one may thus have nitrogen equilibrium without having carbon equilibrium" (p.188). Likewise, his continual stress on the primacy of energy intake and the importance of maintaining this above all else when food supplies are restricted would, if it had been heeded, have saved much confusion in practical nutrition, and many lives, over the past 40 years.

The view expressed in Claude Bernard's manifesto was of course correct in that it suggested the direction in which physiological research was to go; and, indeed, this biochemical approach was to reveal the nature of metabolism. But did we abandon the classical approach to nutrition unnecessarily? Do not the mere statistics of input and output compiled by Lusk paint an equally valid picture of nutrient metabolism as that we see through biochemistry? It remains at present a separate and temporarily irreconcilable picture, as thermodynamics was once from the kinetic theory of gases. But it remains valid. While biochemistry has still to make quantitative statements about nutrient requirements, nutrient balance is still the *sine qua non* of

an adequate diet, the touchstone by which we define a minimal need. And as Lusk shows, page after relentless page, this holistic approach has a coherence of its own — it is possible, by describing the relationship between nutrient input and output, meaningfully to describe the relationship between nutrient needs and environment or physiological state.

So, I hope, I have not reviewed this book as a historical work. I have already found it a valuable text, and a unique source book. It will not find its way in my library to the shelves where books of the same vintage reside, but will be kept much closer to hand.

John Rivers is a Lecturer in the Department of Human Nutrition at the London School of Hygiene and Tropical Medicine, University of London.

Early problems in European economics

Robert Foley

Hunter-Gatherer Economy in Prehistory: A European Perspective.

Edited by Geoff Bailey.

Cambridge University Press: 1983.

Pp.247. £25, \$49.50.

STUDENTS of the European palaeolithic have long required a strong constitution. Years of subsisting on a diet of lithic typologies, soliflucting caves and oscillating glaciers, lightened only by brief excursions into the aesthetics of cave art and the constant comings and goings of obscure culture groups, have imposed a stoicism upon them. With the appearance of books such as *Hunter-Gatherer Economy in Prehistory*, this diet of Würms would appear to be coming to an end.

The past decade has seen the European palaeolithic relinquish its primacy in hunter-gatherer studies to Africa and Australia, and to ethnoarchaeology. This volume of collected papers is intended to return the archaeological wealth of Europe to the mainstream of the discipline. To this end four broad subjects are discussed: theoretical and methodological problems arising from the character of archaeological data; the spatial organization of economic activity; long-term patterns of demography and environment; and large-scale social processes. Editorials by Geoff Bailey link the sections, and Martin Wobst provides a concluding chapter. In substance the volume is not comprehensive. Geographically the south and west dominate at the expense of the rest of Europe, while chronologically the twenty chapters are confined to the period after the appearance of anatomically modern man.

Now in paperback

Science in Action

J Lenihan & J B Fleming (illus)

John Lenihan shares his irrepressible enthusiasm for science, without jargon or technical complexities, but with characteristic wit and humour, while John Fleming's cartoons add a wry and amusing comment to the text.

The whole development of science and technology are presented in such an enjoyable palatable way that one might ask why such a book has not appeared before. New Scientist

June 1983 224pp paperback
0-85498-041-5 £4.95

The Origin and Evolution of Planetary Atmospheres

A Henderson-Sellers
(University of Liverpool)

A fascinating survey of our current understanding of atmospheric evolution within the solar system, reconstructing how very different atmospheres have developed from essentially the same ingredients. This important book is the first to summarise the vast amount of scientific information gathered during the past two decades from direct observation of planetary atmospheres, offering valuable insights into the history of the planets themselves.

July 1983 263pp hardcover
0-85274-385-8 £19.50

Microcomputer Quantum Mechanics

J Killingbeck
(University of Hull)

The increase in availability of microcomputers has led to many students knowing a computer language but not having had the experience of applying it to real scientific problems. Dr Killingbeck seeks to correct this by showing how computation can be integrated with theoretical analysis as part of a unified attack on problems in one of the most interesting areas of modern physics.

April 1983 177pp hardcover
0-85274-445-2 £11.95

Adam Hilger Ltd

Techno House, Redcliffe Way
Bristol BS1 6NX, England
Tel: (0272) 276693 Telex: 449149



The book is a contribution to the literature of the "Cambridge school of palaeoeconomy" founded by Eric Higgs, and in their paper Bailey *et al.* suggest that this school is characterized by three distinctive concepts: site catchment analysis, close man-animal relationships and optimum resource exploitation systems. These are important components of many of the analyses here. What perhaps adds further interest is the attempt to integrate palaeoeconomy with the other developments in hunter-gatherer studies. Thus absent but looming large is the figure of Lewis Binford, and his model of hunter-gatherer variability. Binford contrasted, on the one hand, logistically organized hunter-gatherers (whose subsistence depends on planned and highly structured foraging trips), and on the other, collectors (who subsist more opportunistically in a less-differentiated environment). He was able to show that variability in adaptive strategy had important behavioural and archaeological correlates. The juxtaposition in these papers, especially those of Torrence and of Davidson, of discussion of this model and of the principles of palaeoeconomy highlights many of the difficulties in palaeolithic studies and the extent to which palaeoeconomy has become divorced from other developments in the discipline. For example, although the concept of optimality is employed by several authors little attempt is made either to relate it to its evolutionary basis or to use the techniques of optimal foraging theory. Another example would be the general disregard for the character of archaeological formation processes and the growing field of taphonomy and middle-range theory, such that Bahn (p. 169) can still say that he is "clinging to the hope that faunal remains . . . are representative of the prehistoric economic strategies despite differential preservation and sampling".

These apparent shortcomings are, though, the product of two very positive elements that lie at the heart of this stimulating book. First, it attempts to integrate two traditions in palaeolithic archaeology — the small scale, ethnographically-inspired and behaviourally-based American approach, and the long-term, ecologically orientated British school. And secondly, by presenting a view of the European palaeolithic in which low density populations employ a series of adaptive techniques (ranging from information exchange and alliances to mobility and highly organized hunting strategies) to solve the problems of survival faced by hunter-gatherers in the European glacial environment, these authors have continued the transformation of our perception of palaeolithic peoples from tool-making automatons to real social and biological organisms. []

Robert Foley is a Lecturer in the Department of Anthropology at the University of Durham.

Talking science

Eugene Garfield

The Foreign Language Barrier: Problems in Scientific Communication.

By J.A. Large.

André Deutsch: 1983. Pp.196. £9.95.

FEW would argue with the claim that English is the lingua franca of international science. Therefore, if your native language is not English you face a problem. Unless you can read English you won't know most of what is reported in the literature; and unless you can write and publish in English, your own research may be overlooked by the world scientific community.

J.A. Large, however, examines the language problem from the perspective of scientists whose native language is English. He suggests that researchers in Britain and the United States have been under no pressure to acquire and maintain proficiency in a foreign language, presumably because English has been the dominant language of science for decades. As a result, those speaking and reading only English are ignorant of significant results reported in foreign-language publications.

While the percentage of the world's scientific publications that are published in non-English languages is relatively small, the absolute number is growing. In particular, Japanese and Russian language materials have increased significantly. But it is not clear how much is not covered by the leading abstracting services, which claim to be comprehensive. In any case, the assumption here is that Western scientists are becoming less aware of important research buried in the growing mass of non-English language publications.

Large supports his claims with an array of surveys and studies on library usage of foreign-language materials, world output of publications in various languages and citations to foreign-language publications in journals. His presentation would, however, have been improved if more data had been provided. Each study should be described in terms of the size of the sample and the years to which the data apply. A few additional lines in the text and tables would have allowed the reader to judge how representative and relevant the data are with a minimum of effort.

Although Large's studies and surveys examine in detail the languages of published articles, they do not consider the nationalities of the authors. For example, we are told that English-language articles cite other English-language publications almost exclusively, and only a small proportion of the references cite foreign-language material. But what percentage of the cited English-language items were written by French, German, Russian, Japanese or Chinese authors? Conversely, a significant proportion of references in non-English language articles are to

English-language publications — but how many of these English-language articles were written by scientists from the citing author's country? If scientists in country A cite research conducted in many different nations and reported in English, while scientists in country B only cite papers by their own nation's authors reported in several languages, who is less aware of the international literature? Until we analyse the nationality of the citing and cited authors, in addition to the language of the citing and cited articles, it is premature to characterize the language problem in world scientific communication as a "barrier". We need a comparative bibliometric method that takes into account both interlingual and international links in the literature before we can conclude that there is a language-based crisis in science.

I cannot agree with Large that the language problem currently poses the *biggest* obstacle to scientific communication. In my opinion, the main problem today is information overload. The volume of scientific publication is still increasing, exponentially in some fields. It is more difficult for scientists everywhere to digest everything that is significant in English, much less the foreign-science press. Can we really expect them all to learn Russian, Japanese, Chinese as well as French and German?

Large concludes that computer-based translations will help to solve the language problem. However, indiscriminate translations will only increase the problem of information overload. Information services and review journals have made solid advances in identifying the more significant, high-impact research within the mass of scientific literature. More could be done to identify core material in foreign journals, but only the best should be added to the already overloaded communication channels. The best guarantee that this will occur is through personal contacts between scientists. A comprehensive translation programme would only increase the isolation of scientists by discouraging them from acquiring even a minimal proficiency in a foreign language. If there had not been a cover-to-cover translation programme maybe even more Russian scientists would be publishing in English today. Translation programmes might also decrease attendance at international conferences, where much current scientific information is exchanged.

The cultural and political value of linguistic training is indeed vital to good science. It needs to be promoted not because it will help us deal with the literature better, but rather because it will increase the kind of personal contacts that lead to better identification of important information. []

Eugene Garfield is President of the Institute for Scientific Information, Philadelphia, the publisher of Current Contents and the Science Citation Index.

Tissue culture — use or abuse?

T.M. Dexter reviews the techniques and resources currently available for tissue culture and discusses their future uses.

MOST mammalian cells can now be grown *in vitro* for periods ranging from several days to many years. The period of growth depends on whether the cells are normal or transformed, their stage of development (embryonic versus adult), the tissue, the cell lineage and the species. The increase in culture systems available (both cell lines and primary cultures) in the last twenty or thirty years and the relative ease with which cells can be cultured in quantities sufficient to enable biological, biochemical and molecular studies to be performed, means that more and more people are using tissue culture as a means to an end. A corollary of this is that more people are also abusing tissue culture, paying little or no attention to the physiological significance of the cells that they are growing or whether the cells are normal or not, in their search for the ever elusive "citation classic" on growth regulation, oncogenesis or gene expression!

Objectives

For cell biologists, the objectives of tissue culture are fairly simply defined. Growth, differentiation and development of normal cells are complex processes, requiring both non-specific and specific nutritional and regulatory molecules. One of the aims of tissue culture is to try and dissect these requirements, examining each and every variable in a controllable manner, free of systemic influences. This involves using "pure" cell populations, preferably clones, *in vitro*. Comparisons can then be made with cells undergoing aberrant developmental programmes such as during tumorigenesis.

A first step, of course, was to define the 'non-specific' requirements essential for the survival and growth of many different cell types, such as temperature, osmolarity and pH of the medium, the essential amino acids, vitamins, carbohydrates and trace elements. This work gave rise to the recipes now used for the synthetic growth media supplied by many companies. That the basic formulations of these media have remained unchanged for many years is a reflection of the pioneering efforts in this field by Eagle, Ham, Dulbecco, Fischer and numerous others. However, the successful growth of cells in these media invariably requires the presence of serum — a fickle cocktail of components.

Protein free media

Further elaboration of these synthetic media has led to the development of protein-free media which will support the growth of several different permanent cell lines¹. While this work was of excellence in its approach to defining the nutritional requirements of cells, it was disappointing in that while transformed cells grew quite happily in such synthetic media for extended periods, non-transformed (normal) cells still generally required the presence of serum or serum components. (By transformed cells, I mean cells which were grown from normal tissue and became progressively adapted for growth in tissue culture, or cells which were taken from tumour tissue and showed an ability to grow *in vitro*. Invariably, such cells are aneuploid and have been selected for their ability to grow autonomously *in vitro* with minimal nutritional supplementation. Whether the cells are tumorigenic is another matter! By normal cells, I mean primary explants or continuously growing cells which maintain a diploid karyotype, a consistent phenotype and are non-tumorigenic. In only relatively few instances has it been determined if these cultured cells can function *in vivo* like the corresponding normal cells).

Following these efforts, extensive work has gone into defining the serum components essential for the maintenance of normal cells *in vitro*. These include corticosteroids, insulin, thyroid and pituitary hormones, various growth factors, "spreading" factors and transferrin². The mode of action of many of these supplements, and the nature of the cell responses mediated by them, are still largely unknown. Interestingly, however, the transferrin receptor has recently been implicated in cell proliferation³ and transferrin has been found to be an essential component for the serum-free growth of many cell types, including haemopoietic progenitor cells⁴. From such humble beginnings is biology made.

Specific versus non-specific

Like the essential amino acids and trace elements, I suspect that many of these other serum agents, such as transferrin, insulin and the steroid hormones, will be "non-specific" for growth. Instead they will have a tissue and lineage restriction

only in terms of the quantities of the elements required. Rather more exciting are the current studies being undertaken with the so-called growth factors. These include nerve growth factor (NGF), epidermal and fibroblast growth factors (EGF and FGF), angiogenesis factor and various haemopoietic cell growth factors — many with a marked tissue or cell lineage restriction in their action. This is a topical and rapidly expanding field, too vast to be adequately touched on here, and the reader is referred to a good review on these factors⁵.

The importance of these growth factors (many of them commercially available) is two fold. Many if not all normal cells demonstrate an absolute dependence on these growth factors for survival and proliferation. Commercially available sera may either not contain the growth factors, or contain them in too low a concentration to permit growth. Also, normal cells are usually grown in xenogeneic-sera; for example, human cells are grown in fetal calf serum and murine cells in horse serum. Some of these growth factors, such as the haemopoietic cell growth factors are species specific in their activity. Thus, the inability to grow many normal cells may be due to a deficiency or an absence of the appropriate growth factor. Furthermore, the growth of 'transformed' cells in the absence of serum (and of these growth factors) may represent a fundamental process in the progression toward malignancy. Investigations are now underway to determine whether this independence arises as a result of the autonomous production of the essential growth factors by the tumour cells.

Thus, tissue culture is entering a new era of growth, although obviously many questions and problems remain. Also apparent is a realization of the limitations imposed by tissue culture, and of its prime requirement, namely that the culture cells must reflect their origin. □

1. Katsuta, H. & Takaoka, T. *Meth. Cell Biol.* **6**, 1-42 (1973).
2. Barnes, D. & Sato, G. *Rev. analyt. Biochem.* **102**, 225-270 (1980).
3. Trowbridge, I.S. & Omary, M.B. *Proc. natn. Acad. Sci., U.S.A.* **78**, 3039-3043 (1981).
4. Guilbert, L.G. & Iscove, N.N. *Nature* **263**, 594-595 (1976).
5. Baserga, R. (ed.) *Tissue Growth Factors*. (Springer, Berlin, 1981).

T.M. Dexter is head of experimental haematology at the Christie Hospital and Holt Radium Institute, Manchester, U.K.

General equipment for tissue culture

● A new MicroWell plate with a conical bottom has been added to the range of Nunc plates. The new MicroWell plate is primarily intended for use in serology and bacteriology and is available in two versions, either standard hygienic or sterilized by irradiation. The plate is of standard dimensions and fits into existing ancillary equipment. The V-shaped MicroWell plate is made of polystyrene, and the wells have a capacity of 300 μ l.

Circle No. 100 on Reader Service Card.

● Dynatech Laboratories have extended their range of tissue-culture-treated micro-test plates to include a tissue typing design. This can be used for fluorescent homologous leucocytic antibody (HLA) and standard dye cytotoxicity tissue typing tests and other microassay techniques such as virus neutralization and cell cloning. The plates come in 3 different configurations, each with the same external plate dimensions — a 60-well Terasaki format plus 72 and 120 well versions. In all three, the well is surrounded by a raised shoulder to prevent cross-contamination. The plates are compatible with standard multi-gang syringes such as the 'Hamilton' type.

Circle No. 101 on Reader Service Card.

● The Camlab glassblowing service supply roller bottles and tubes for tissue culture. The roller bottles are available in two lengths — 33 cm or 50 cm tall — with round or sloping shoulders. Body diameter is 100 mm, although other sizes can be supplied. The tissue culture tubes consist of Pyrex test tubes (150 $\times$ 16 mm) having one or two flat windows each measuring 35 $\times$ 8 $\times$ 15 mm. The tubes are supplied plain or with plastic screw caps and rubber discs and liners.

Circle No. 102 on Reader Service Card.



The Camlab roller bottles

● Nunc's six-well multidish with special surface treatment is designed to maximize cell adherence and growth. Moulded from Nunclon Delta, the six numbered wells are separated by reservoirs which will reduce cross contamination or will hold sterile water to reduce the evaporation of media.

Circle No. 103 on Reader Service Card.

● The Dilvac collection of tissue culture roller bottles is now available through Day-Impex. The bottles are made from borosilicate glass which reduces imperfections in the glass surface and ensures greater cell yields and more uniformity of growth. The bottles are available with autoclavable screw caps and with either wide or narrow necks.

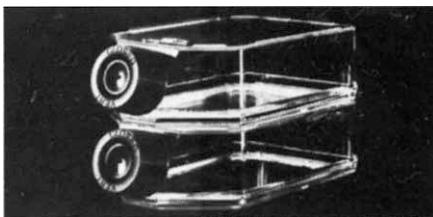
Circle No. 104 on Reader Service Card.

● The new design of the Nunc 260 ml tissue culture flask permits uniform free air passage. Designed to allow access to the cell monolayer, the flask has a short, wide angled neck. The two-position cap marks the correct position for venting or closure. The flask is made from Nunclon Delta, which has been selected to increase cell growth and optical clarity.

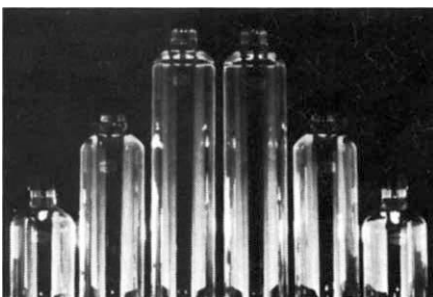
Circle No. 105 on Reader Service Card.

● Drummond Scientific has recently introduced the new model 390 multiple or fixed volume repeating micropipette. The pipette is designed to deliver a maximum volume of 1,000 μ l, and four spacers are provided which will adjust the unit to deliver 100 μ l, 250 μ l, 500 μ l, and 800 μ l. Custom spacers to accommodate other volumes are also available. The pipette is supplied with a repeating kit, 4 standard volume spacers, 7 extra barrels, 2 plungers and a maintenance kit.

Circle No. 106 on Reader Service Card.



Nunc's 260ml flask



Dilvac's roller bottles

● A new tissue culture chamber is now available from Medical Systems. The LU-CB-1 chamber facilitates the observation of cell motility under a microscope and aids the electrophysiological recording of membrane potentials from single cells. The culture medium is covered with mineral oil which prevents evaporation and enables an air/CO₂ mixture to permeate the medium to ensure constant pH.

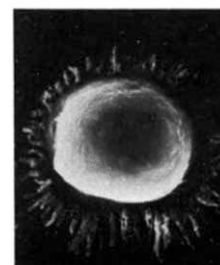
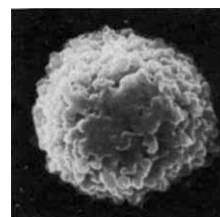
Circle No. 107 on Reader Service Card.

● A new range of culture dishes is now available from Zer Hitec. These are coated with extracellular matrix (ECM), a basement-membrane-like substance which enhances cell adherence and proliferation. The ECM is produced by allowing endothelial cells to grow just above the plastic surface of the dish. These cells then secrete the ECM and, when the ECM is complete, are removed leaving the matrix intact. The organization and chemical composition of the ECM is similar to basal laminas occurring *in vivo*. Unlike reconstituted matrices, naturally produced ECM has its various components in their native configuration and proportion. Cells plated on ECM exhibit higher plating and cloning efficiencies, have longer life-spans, reach a higher saturation density and have lower requirements for serum and added growth factors. They also respond better to physiologically occurring hormones and growth factors, express differentiated properties and attach and migrate rapidly. Plating on ECM provides a means to develop new cell lines.

Circle No. 108 on Reader Service Card.

● Life Science Laboratories have recently introduced the Sonabox, a multi-purpose noise reduction cabinet for use with laboratory ultrasonic cell disruptors. Constructed from cyanamid quiet metal, the Sonabox is insulated to provide a noise reduction of 50 db.

Circle No. 109 on Reader Service Card.



Comparing a megakaryocyte maintained for 24 hours on plastic (top) with one maintained on Zer Hitec's ECM (bottom).

● The new beaded agarose gel from Zer Hitec is designed for use in affinity chromatography. Substances bind to the matrix through covalent bonding to the beads of specific biochemical reagents. Prepared from a 4% solution of purified agarose, the gel is composed of spherical beads (50–150 μm in diameter) suspended in water with a preservative to prevent bacterial growth. In the form supplied, the beads are stable at room temperature indefinitely. This stability remains unimpaired over a variety of conditions of temperature, pH and molarity. Rough freezing should be avoided, and prolonged temperatures above 40°C may lead to irreversible damage of the bead structure. Agarose beads are also available with pre-bound immunoglobulins, enzymes and active proteins. Although dependent on conditions, 1 ml of Zer Hitec pre-bound goat-anti-rabbit immunoglobulin (IgG) will generally bind 3 mg of rabbit IgG. These beads are recommended for use in routine radioimmunoassay techniques diluted with plain agarose beads.

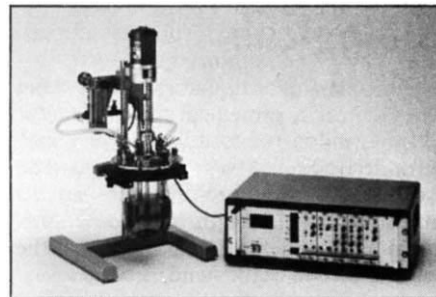
Circle No. 110 on Reader Service Card.



Life Science Laboratories' sonicator

● A range of sonicators, or ultrasonic probe disruptors, is now available from Life Science Laboratories. It covers power outputs from 10 to 2,500 W and handling volumes from a few millilitres to several hundred litres. As well as disrupting cell walls, which permits the extraction of sub-cellular components, these units may be used as homogenizers, emulsifiers, solubilizers or depolymerizers, for disaggregation, or even to stimulate reactions to take place.

Circle No. 111 on Reader Service Card.



Astra Scientific's 2 litre fermenter

● A new, automatic laboratory fermenter with a 2 litre capacity culture vessel is now available from Astra Scientific. The microprocessor-controlled unit has a system of rack-mounted, plug-in modules for regulation and measurement. Any control parameter can be called up on the LED display, or the unit can be set to automatically scan a group of measurements. The biological vessel is borosilicate glass and will withstand a pressure of 30 p.s.i. The unit can be dismantled from its frame.

Circle No. 112 on Reader Service Card.

Incubators

● Boro Labs have recently launched the dual chamber version of their Salamander CO₂ incubator. This effectively gives two separate incubators in the same chassis or cabinet, each with their own heaters, controls, CO₂ supply and specific options. This allows comparative investigations to be carried out simultaneously and therefore ensures the minimum variation in conditions. The incubator incorporates a water jacket on five sides of the chambers to increase temperature stability. A selection of options is also available, including a direct moisture injection relative humidity control, an automatic cylinder change-over system and cooling coils.

Circle No. 113 on Reader Service Card.

● Available from Heraeus are two low temperature incubators, the BK 5060 and the BK 500. They are suitable for the storage and treatment of samples at temperatures above 0°C. Examples of their applications include the storage of solutions, sera and culture plates before use; the incubation of bacterial and fungal cultures at 20 to 25°C and the typing of tissues in the preparation of transplants. Also available is the Heraeus growth chamber (BK 5060 EL) which is equipped with two electronic controllers, for day and night operation, and with a time switch for the preselection of day and night programs.

Circle No. 114 on Reader Service Card.

● Genetic Research Instrumentation now have a series of new incubators including automatic CO₂ incubators, plant cell culture cabinets, general purpose incubators and incubator rooms. The new shaking incubator has a maximum capacity of twelve 2,000 ml flasks with a temperature range of 0 to 75°C and a maximum shaking speed of 4,000 r.p.m. The incubator is available with integral lights and cycling timers, a refrigeration system for below ambient operation and flask holders and platforms.

Circle No. 115 on Reader Service Card.

● In addition to the existing roller culture apparatus available from New Brunswick Scientific there are now several new products in the tissue culture field. A range of four new CO₂ incubators has recently been launched; two large capacity forced warm air incubators and two smaller capacity water jacketed incubators. These incubators offer very accurate control of temperature, humidity and CO₂ with a recorder output for all three parameters. Also available is a blast freezer capable of fast-freezing large amounts of material to low temperatures without involving solid CO₂, alcohol or liquid nitrogen. The blast freezer delivers an air stream at -80°C at a very high speed to the material to be frozen. Two models are available, the larger having a capacity of 70 litres and the smaller one 40 litres. The third new product is a controlled-rate freezer which uses a cascaded refrigeration system to achieve -75°C. The operator can program the rate freezer by entering four temperature ramps, each having its own hold and slope facility. Slopes can be set from 0.1°C per minute to 1.5°C per minute.

Circle No. 116 on Reader Service Card.

● A range of biological incubators is now being produced by Astell Hearson. There are three incubators of capacities 91 litres, 136 litres and 200 litres. The aluminium chambers are heated by mat heaters. Temperature is controlled by a solid state proportional control thermostat and all the incubators are fitted with a safety thermostat and an inner glass door, to which is attached a mercury in glass check thermometer.

Circle No. 117 on Reader Service Card.

● Two new mechanical convection incubators are now being supplied by National Appliance Company. The new incubators, designated models 301 and 302, are intended for small unit applications. The 302 model has CO₂ capabilities. The chamber interiors of both models are made of stainless steel and snap out. The impellers for the chamber atmosphere convection system are disposable, as are the air and CO₂ filters. All the electrical wiring system and the components are accessible from the top of the units through a removable cover. In both the 301 and 302 the automatic temperature controls are electronic, and the true chamber temperatures are indicated by a built-in dial thermometer. A safety thermostat provides heat regulation. By using a sealed chamber and mechanical convection, these incubator units reduce CO₂ flow rates without stratification. Chamber integrity is aided by glass inner doors and special vapour-lock gaskets, which reduce atmosphere loss during routine inspections. The 302 model has a built-in CO₂ flowmeter, which provides a constant prescribed mixture of CO₂ and air, and a CO₂ purge timer restores CO₂ tension.

Circle No. 118 on Reader Service Card.

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.

Biologicals

● Fibroblast growth factor (FGF) from Zer Hitec is a growth promoting factor which is mitogenic to a variety of mesoderm-derived cells at nanogram levels. The use of FGF in cultures permits *in vitro* cultivation of cells with fastidious nutritional requirements. It also lowers the dependence on serum and prolongs survival and the replicative life-span. FGF also accelerates cell proliferation, increases the colony-forming and plating efficiency by allowing cells to proliferate at clonal cell densities and permits expression of differentiated characteristics. Zer Hitec FGF has a molecular weight of 13,000 (isoelectric point 9.6). It is isolated from bovine brain and is homologous in structure to a fragment of bovine myelin basic protein. It is supplied as a purified, water soluble, lyophilized powder.

Circle No. 119 on Reader Service Card.

● Lympho/Scan anti IgM is a new immunoperoxidase staining kit that has been introduced by Biomeda for the detection of human μ heavy chain in histological sections of lymphatic and other tissues. The identification of the distribution of cytoplasmic and cell surface μ heavy chain in lymphoid tissues is valuable in the differentiation of B-cell proliferations. This new immunohistochemical human IgM kit is part of a series of seven Lympho/Scan immunoperoxidase staining kits to be released soon by Biomeda. The procedure employed in this kit is based on the avidin-biotin affinity cytochemistry. It eliminates the need for secondary and linking antibodies so that reagent incubations can be reduced to 5–10 minutes for the IgM specific antibody followed by two 5 minutes incubations for the peroxidase reagent and the chromogen developer.

Circle No. 120 on Reader Service Card.

● New from Sera-lab is the Insolmer kit which renders antigens insoluble. The kit consists of a 5 ml vial of a fine suspension of polyester particles, a 2.5 ml vial of polymer-activating solution and a 2.5 ml vial of a cross-linking reagent.

Circle No. 121 on Reader Service Card.

ADVERTISEMENT

Thermocouple THERMOMETERS

We make 14 thermometers, with digital or meter readout. All read in 2 secs, are extremely easy to use. Simply apply probe — read temperature.

Unbreakable probes are steel sheathed including World's Smallest Microprobe, 0.01" dia.

Models

BAT-4 BIOLOGICAL

BAT-5 CRYOGENIC

BAT-12 DIGITAL

BAT-8 DIGITAL °C or °F

TH 3 NEW CLINICAL TYPES



BAILEY INSTRUMENTS INC.
Saddle Brook, N.J. 07662 U.S.A.

Circle No. 25 on Reader Service Card.

● CNBr-epidermal growth factor (CNBr-EGF) is now being produced by Biomedical Technologies. CNBr-EGF is a biologically inactive EGF analogue which retains its receptor binding activity. It has been shown that CNBr-EGF treated cells in culture can have a full restoration of biological activity by the addition of anti-EGF antibodies, and so CNBr-EGF can be used for the investigation of the EGF receptor mediated response.

Circle No. 122 on Reader Service Card.

● A new, fully defined, low protein, serum-free culture medium is now available from Boehringer Mannheim Biochemicals. It consists of the Iscove modification of minimum essential medium (MEM) Dulbecco, plus three additives: bovine serum albumin (BSA), transferrin and soybean lipids. This system has been used to culture hybridoma cell lines, T-lymphocytes, B-lymphocytes and macrophage cells.

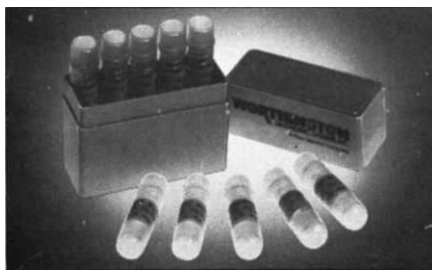
Circle No. 123 on Reader Service Card.

● T-marker quality control sera have been developed by NMS Pharmaceuticals. The controls consist of high and low reference values which can be used with all the current tumour marker RIA kits available. The controls were assayed through Medex Laboratories, the clinical division of NMS Pharmaceuticals. Included in the assay ranges are values for: ACTH, AFP, aldosterone, CEA, calcitonin, CK-B, HCG, ferritin, gastrin, PAP, prolactin, PTH, and thyroglobulin. The product consists of lyophilized serum base and is packaged in four 2 ml vials per box.

Circle No. 124 on Reader Service Card.

● A purified ricin-A chain that is free of the ricin-B chain transport agent and whole toxin has been announced by Worthington Diagnostics Systems. It is chemically stable in solution for extended periods and is 10,000 times less toxic to whole cells than complete ricin. It is now being made available to research hospitals, universities, and medical and government research laboratories. Worthington ricin-A chain remains stable in solution for 6–8 months when stored at -70°C . The product can be conjugated to monoclonal antibodies, hormones or other ligands to target cytotoxic effect, and is suitable for injection into whole animals at the concentrations required for targeted cytotoxic experiments.

Circle No. 125 on Reader Service Card.



The ricin-A chain

● A new zinc chelate affinity adsorbent is now available from Boehringer Mannheim Biochemicals. The adsorbent consists of chelated zinc bound to agarose through a spacer. The adsorbent will bind any protein which forms a complex with zinc, especially the general protease inhibitor α_2 -macroglobulin. When α_2 -macroglobulin is bound, the column can be used to remove endoproteases from any solution of interest.

Circle No. 126 on Reader Service Card.

● Three kinds of diagnostic pneumococcal antisera are available from Uniscience. They are intended for the identification and typing of pneumococci by means of a capsular reaction test and the counter-current immunoelectrophoresis of clinical specimens. The three sera are produced by Statens Serum Institut. The first, Omni-Serum, contains antibodies to all known pneumococcal capsular and polysaccharide-type antigens. Nine pooled sera each react with seven to eleven types of pneumococci, together covering all 83 types. The last is 46-type or Group Sera, which reacts with single types or with groups numbered 1 to 48.

Circle No. 127 on Reader Service Card.

● Gibco has extended its range of tissue culture products with the addition of two new media. The 309 fetal calf serum is aseptically collected to ensure a low endotoxin level. This medium will support those myelomas currently used to produce cell hybrids. Gibco is also producing a chemically defined medium — Iscove's medium — along with its supplements. This medium was developed primarily for lymphocyte culture, but has since been found to support other cell lines.

Circle No. 128 on Reader Service Card.

● The new Institute test kits from Zer Hitec are designed for radioimmunoassay (RIA) testing procedures. They remove the need for centrifugation, decanting and the addition of precipitating reagents as follows. The sample, antiserum and labelled hormone are pipetted into the Institute. After the first incubation, the stopper is removed and the contents allowed to drain. The tube is then prepared for counting by washing after the second incubation. Institute test kits are available for β -HCG, LH, FSH, prolactin, T_2 , T_4 , TSH and 17β -estradiol.

Circle No. 129 on Reader Service Card.



Zer Hitec's RIA kits